

AD _____
(Leave blank)

Award Number: DAMD17-03-1-0251

TITLE: Rational Inhibitors of DNA Base Excision Repair Enzymes:
New Tools for Elucidating the Role of BER in Cancer
Chemotherapy

PRINCIPAL INVESTIGATOR: Daniel J. Krosky

CONTRACTING ORGANIZATION: Johns Hopkins University
Baltimore, MD 21218-2686

REPORT DATE: November 2006

TYPE OF REPORT: Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

☒ Approved for public release; distribution unlimited

☐ Distribution limited to U.S. Government agencies only;
report contains proprietary information

The views, opinions and/or findings contained in this report are
those of the author(s) and should not be construed as an official

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</small>					
1. REPORT DATE (DD-MM-YYYY) 21-11-2006		2. REPORT TYPE Addendum		3. DATES COVERED (From - To) 21 APR 2003 - 20 OCT 2006	
4. TITLE AND SUBTITLE Rational Inhibitors of DNA Base Excision Repair Enzymes: New Tools for Elucidating the Role of BER in Cancer Chemotherapy				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER DAMD17-03-1-0251	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Daniel J. Krosky, P. I.				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Johns Hopkins University Homewood Research Admin 105 Ames Hall 3400 North Charles Street Baltimore, MD 21218-2686				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The aim of this research program was to investigate the role of the uracil base excision repair (UBER) pathway in the mechanism of action of the antineoplastic agent 5-fluorouracil (5-FU) through the discovery and development of inhibitors against UBER enzymes. Over the course of this funding period, we have developed novel, potent oligonucleotide and small molecule inhibitors of the first enzyme in the UBER pathway, uracil DNA glycosylase (UDG) as outlined in Tasks 1 and 2 of the approved Statement of Work. Using these inhibitors, we were able to greatly expand our understanding of how UDG recognizes and interacts with normal and uracil-containing DNA. However, the limited cell penetrance and/or potencies of these compounds hindered their use in dissecting the interactions between UDG and 5-FU in cellulo (Task 3 of the approved Statement of Work). Instead, a yeast genetics approach was used, where the cytotoxicity and potency of 5-FU was determined against a panel of UBER mutants. Consistent with the model that UBER potentiates the cytotoxicity of 5-FU through futile DNA repair, a UDG knockout yeast strain is resistant to 5-FU, while a yeast strain that lacks the ability to process abasic sites (a UBER repair intermediate) is hypersensitive to 5-FU. These results suggest that inhibitors of abasic site processing enzymes in the UBER pathway could be useful agents in the potentiation of 5-FU antineoplastic activity.					
15. SUBJECT TERMS Uracil DNA glycosylase, inhibitors, 5-fluorouracil, breast cancer					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Unlimited	18. NUMBER OF PAGES 52	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code)

Table of Contents

Introduction.....	3
Body.....	4
Key Research Accomplishments.....	6
Reportable Outcomes.....	7
Conclusions.....	7
References.....	7
Appendices.....	5 publications

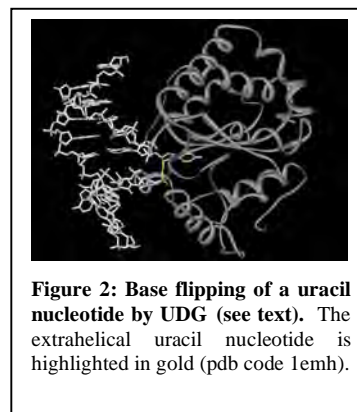
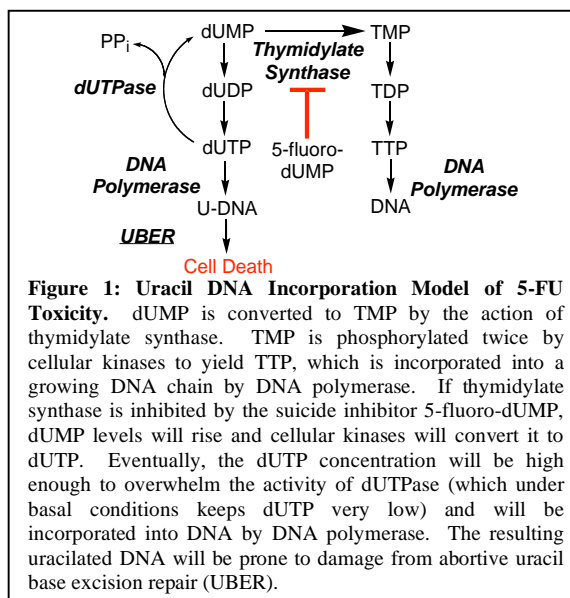
INTRODUCTION

This program seeks to obtain a fundamental understanding of the chemical mechanisms by which enzymes repair damaged DNA, and to use this information to design small molecule inhibitors of these enzymes. The driving force for these efforts is the recognition that the effectiveness of cancer chemotherapy regimes is intimately connected to, and in some cases directly relies on, DNA damage repair pathways. A more sophisticated understanding of the roles of DNA damage repair in the pharmacology of DNA replication inhibitors will allow for the design of better treatments against breast and other cancers.

One particularly important chemotherapeutic agent whose mode of action has been proposed to be intimately linked with DNA damage repair pathways is 5-fluorouracil (5-FU) and its congeners, which are used to treat breast cancer, as well as cancers of the colon, head and neck (1, 2). It is well known that 5-FU targets the biosynthetic enzyme thymidylate synthase (TS), which generates thymidine monophosphate from deoxyuridine monophosphate (dUMP) (3). However, the mechanism(s) of how thymidine depletion results in cytotoxicity (“thymineless death”) has

been unclear. One proposed model (Figure 1) suggests that as a result of TS inhibition, cellular dUMP levels begin to rise and is efficiently converted to deoxyuridine triphosphate (dUTP). Most DNA polymerases cannot discriminate between dUTP and TTP, and thus the high dUTP and low TTP levels results in the misincorporation of uracil opposite adenine in replicating DNA. While these uracils are not mutagenic per se, they are nonetheless removed by the uracil base excision repair (UBER) pathway. The high density of uracil in DNA, combined with the low concentration of TTP, lead to futile repair cycles which would eventually result in cell death either through DNA fragmentation and/or the accumulation of toxic repair intermediates.

In order to begin to explore the relationship between the cytotoxic mechanism of 5-FU and uracil base excision repair, we initiated a research program to discover inhibitors to the initial enzyme in the UBER pathway, uracil DNA glycosylase (UDG – Tasks 1 and 2 of the Approved Statement of Work). Once UDG inhibitors were



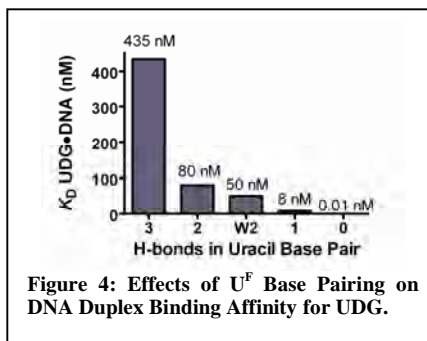
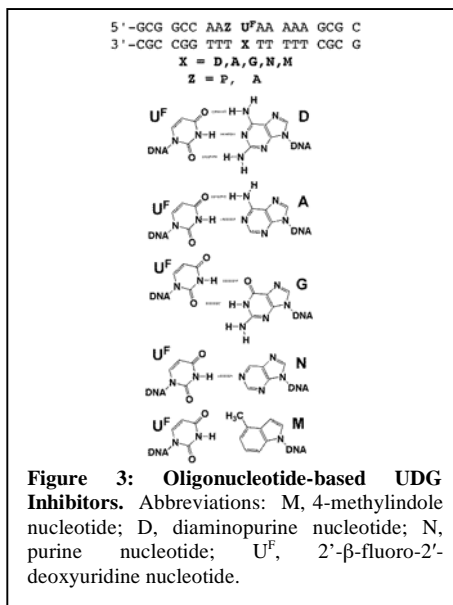
identified, their effects on 5-FU cytotoxicity and potency in cell culture would be evaluated (Task 3 of the Approved Statement of Work).

BODY

Oligonucleotide-based UDG Inhibitors: UDG, like other DNA glycosylases, couples substrate recognition and enzyme catalysis through base flipping, a process where the cognate damaged base is rotated $\sim 180^\circ$ from inside of the DNA double helix and into the enzyme active site. In the case of UDG, only deoxyuridine can attain this fully flipped conformation due to the tight induced fit and high complementarity of uracil within the enzyme active site (Figure 2). Thus, short oligonucleotides containing a non-hydrolyzable deoxyuridine analog (β -2'-fluorodeoxyuridine; U^F) are specific, sub-micromolar UDG inhibitors.

In an effort to both improve potency and develop a set of mechanistic tools for the study of enzymatic base flipping, a series of U^F -containing double stranded oligonucleotides were synthesized where the number of the hydrogen bonds in the U^F base pair was systematically altered (Figure 3). During the course of extruding the target base from within the DNA duplex to the active site, the uracil base pair must be broken. The energetic penalty for disrupting the uracil base pair is paid for by the intrinsic binding energy of the uracil-DNA-UDG complex (i.e. the sum of all of the exergonic interactions). We proposed that the ablation of hydrogen bonds in the uracil base pair would lead to increased affinity of those oligonucleotides with UDG, since less binding energy would be expended by the system to flip the uracil out of the DNA duplex.

As seen in Figure 4, the removal of three hydrogen bonds from the uracil base pair ($U^F \cdot D$ vs. $U^F \cdot M$) results in a 43,000-fold increase in binding affinity ($K_D = 435$ nM and 0.01 nM, respectively). A detailed mechanistic analysis using transient kinetic approaches and linear-free energy relationships suggested that the origins of the enhanced affinity of UDG for $U^F \cdot M$ was due not only to removing the energetic penalty for breaking the base pair hydrogen bonds, but also largely due to the enhanced local flexibility of the DNA. This increase in conformational freedom allows for the DNA to more easily adopt the bent conformation observed in the



fully flipped out complex (Figure 2). These results were published as two articles in *Biochemistry* (see Reportable Outcomes).

Small Molecule UDG Inhibitors: While the potency of the oligonucleotide-based inhibitors was vastly increased by a relatively small chemical manipulation, they were still less than ideal tool molecules for cellular or *in vivo* studies due to their poor drug-like properties. Thus, we initiated a small molecule UDG inhibitor discovery effort. Initially, a high-throughput screening assay was developed, and a commercial 10,000 compound library was screened against human UDG. No authentic human UDG inhibitors were confirmed upon follow-up of the screening hits.

In order to enhance the probability of inhibitor discovery, we developed a novel chemical library focused against UDG. The bias of this library was introduced by taking advantage of the specific interactions the enzyme makes with extrahelical uracil in the active site. UDG only binds uracil with modest affinity ($K_D \sim 80 \mu\text{M}$), but uracil is highly complementary to the active site. Thus, we used uracil as an anchor into the active site of

UDG, and then relied on random exploration of adventitious binding sites proximal to the active site by tethering random small molecules to one of three formyluracils (Figure 5).

The oxime chemistry used in the library construction is ideal for use in high-throughput screening (HTS) because the reactions are quantitative, require no purifications, and the resultant mixtures can be used directly in the HTS assay. In addition, the ability to easily vary the linker length between the formyluracil and the variable aldehyde allows for increased chemical diversity in the library and rapid optimization of the spacing between the uracil and the variable binding element

Using these libraries, an initial hit containing a catechol moiety, 3-(3)-13, was

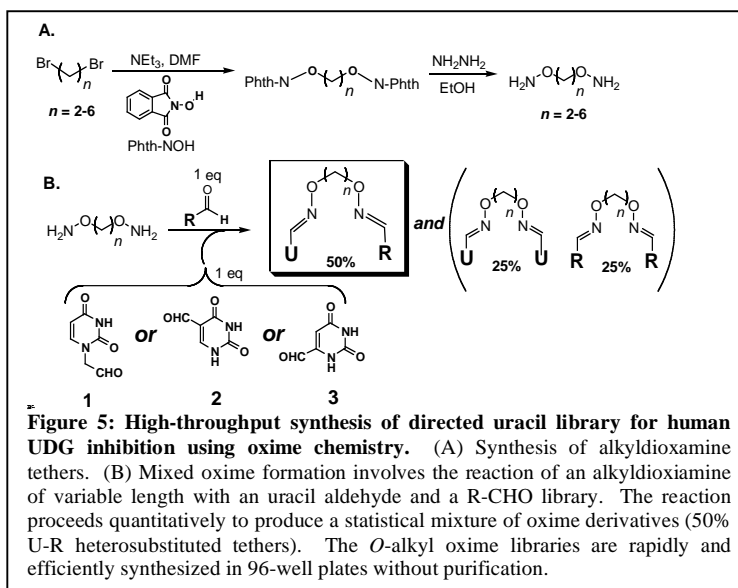
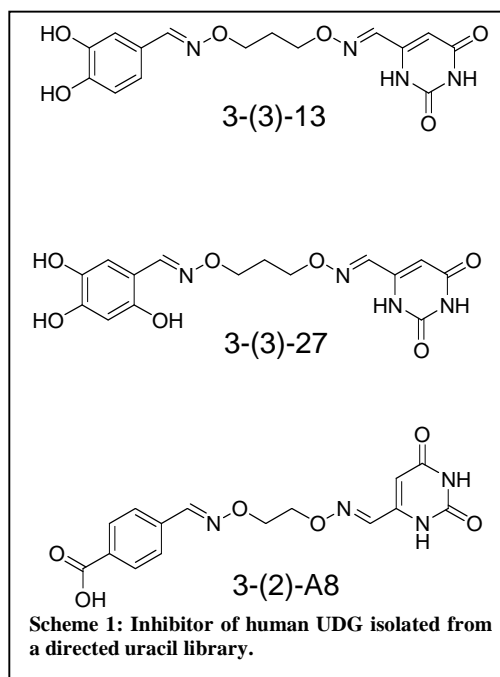
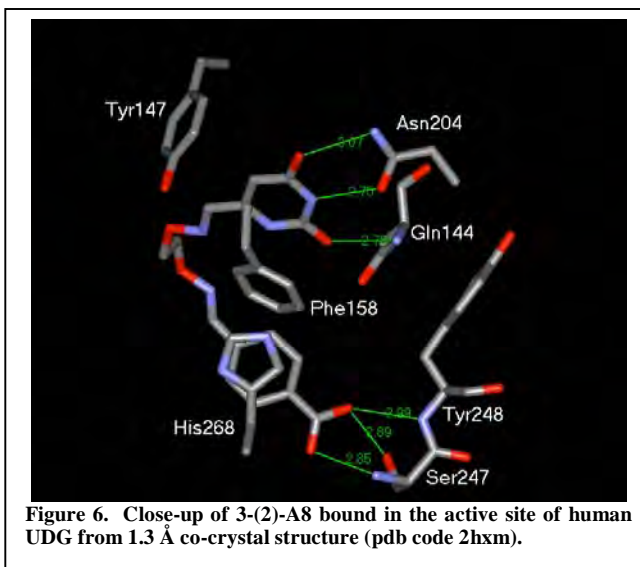


Figure 5: High-throughput synthesis of directed uracil library for human UDG inhibition using oxime chemistry. (A) Synthesis of alkylldioxamine tethers. (B) Mixed oxime formation involves the reaction of an alkyldioxamine of variable length with an uracil aldehyde and a R-CHO library. The reaction proceeds quantitatively to produce a statistical mixture of oxime derivatives (50% U-R heterosubstituted tethers). The *O*-alkyl oxime libraries are rapidly and efficiently synthesized in 96-well plates without purification.



identified as an inhibitor of human UDG (Scheme 1). The affinity of this compound ($K_i = 1.6 \mu\text{M}$) for human UDG was 50-fold greater than uracil. Several analogs of 3-(3)-13 were made, which led to the discovery of an analog, 3-(3)-27 (Scheme 1), with a 5-6-fold increase in potency ($K_i = 0.3 \mu\text{M}$). The chemical instability of these catechol-containing inhibitors toward oxidation limited their usefulness beyond in vitro characterization. However, an analog that replaced the vicinal hydroxyl groups with a carboxylic acid (3-(2)-A8, Scheme 1) was nearly as potent as the initial catechol hit ($K_i = 3.2 \mu\text{M}$) and is chemically stable. An x-ray co-crystal structure of 3-(2)-A8 and human UDG was solved at 1.3 Å resolution, and revealed that the inhibitor mimics the shape and interactions of damaged DNA with the enzyme (Figure 6). Currently, we are using this co-crystal structure as a guide to design more potent analogs of 3-(2)-A8 that will be useful in cell-based studies (see below). These studies resulted in three publications (see Reportable Outcomes).

The current small molecule inhibitors of human UDG were insufficiently potent and/or bioavailable to significantly affect cellular UDG activity. In the absence of suitable tool compounds, the interactions of uracil base excision repair and 5-fluorouracil chemotherapy were elucidated in a yeast model system (Seiple, et al. *Nucleic Acids Res.* **2006**, 34, 140-51). These studies strongly suggested that UDG potentiates the cytotoxic effects of 5-fluorouracil by initiating futile repair of the high concentrations of deoxyuridine in the DNA introduced as a consequence of thymidylate synthase inhibition. Consistent with this model were the observations that UDG knockout yeast were resistant to the effects of 5-FU, while yeast that were deficient in AP endonuclease activities were hypersensitive to 5-FU (*i.e.* a high density of uracil in DNA is not toxic per se, but the attempted repair of these lesions triggers cell death). Studies that extend these observations from the yeast model system to mammalian cells are ongoing.



KEY RESEARCH ACCOMPLISHMENTS

- Developed and characterized highly potent oligonucleotide-based inhibitors of uracil DNA glycosylase (SOW, Task 1 and 2).
- Developed novel chemical libraries based focused against DNA repair enzymes (SOW, Task 1).
- Identified potent and selective inhibitors of human uracil DNA glycosylase (SOW, Task 1 and 2).

- Solved x-ray co-crystal structure of 3-(2)-A8 and human uracil DNA glycosylase at 1.3 Å resolution.

REPORTABLE OUTCOMES

This work has resulted in five published manuscripts:

1. Krosky, DJ, Schwarz, FP, Stivers, JT Linear Free Energy Correlations for Enzymatic Base Flipping: How Do Damaged Base Pairs Facilitate Specific Recognition? *Biochemistry* **2004**, *43*, 4188-4195.
2. Krosky DJ, Song, F, Stivers, JT The Origins of High-Affinity Enzyme Binding to An Extrahelical DNA Base. *Biochemistry* **2005**, *44*, 5949-5959.
3. Jiang YL, Krosky DJ, Seiple LM, Stivers JT. Uracil-directed Ligand Tethering: An Efficient Strategy for Uracil DNA Glycosylase (UNG) Inhibitor Development. *J Am Chem Soc* **2005**, *127*, 17412-20.
4. Jiang YL, Chung S, Krosky DJ, Stivers JT. Synthesis and high-throughput evaluation of triskelion uracil libraries for inhibition of human dUTPase and UNG2. *Bioorg Med Chem* **2006**, *14*, 5666-72.
5. Krosky, DJ, Bianchet, MA, Seiple, L, Chung, S, Amzel, LM, Stivers, JT Mimicking Damaged DNA with a Small Molecule Inhibitor of Human UNG2. *Nucleic Acids Res* **2006**, *34*, 5872-9.

CONCLUSIONS

The long-term goal of this research is to increase the effectiveness of 5-fluorouracil chemotherapy through selective targeting of base excision repair enzymes. A first step to achieving this end is to gain a better understanding of its mechanism of action through the development of specific inhibitors of UBER enzymes to use as tool compounds. We have progressed towards this goal by developing methodology to create chemical libraries focused against DNA repair enzymes, and we have successfully used this technology to discover the first non-nucleotide inhibitors of human UDG. At least one of these inhibitors shows evidence of being cell permeable, and will be used as the basis to design more potent molecules to elucidate the mechanism of action of 5-fluorouracil against breast cancer cells.

REFERENCES

1. Malet-Martino, M. and Martino, R. (2002) Clinical studies of three oral prodrugs of 5-fluorouracil (capecitabine, UFT, S-1): a review. *Oncologist*, **7**, 288-323.
2. Rich, T.A., Shepard, R.C., Mosley, S.T. (2004) Four decades of continuing innovation with fluorouracil: Current and future approaches to fluorouracil chemoradiation therapy. *J. Clin. Oncol.* **22**, 2214-2232.

3. Santi, D.V., McHenry, C.S., Raines, R.T., Ivanetich, K.M. (1987) Kinetics and thermodynamics of the interaction of 5-fluoro-2'-deoxyuridylate with thymidylate synthase. *Biochemistry*, **26**, 8606-8613.

APPENDICES

Five manuscripts are included (see above).

Linear Free Energy Correlations for Enzymatic Base Flipping: How Do Damaged Base Pairs Facilitate Specific Recognition?[†]

Daniel J. Krosky,[‡] Frederick P. Schwarz,[§] and James T. Stivers^{*,‡}

Department of Pharmacology and Molecular Sciences, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, and The Center for Advanced Research in Biotechnology and National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received December 22, 2003; Revised Manuscript Received February 6, 2004

ABSTRACT: To efficiently maintain their genomic integrity, DNA repair glycosylases must exhibit high catalytic specificity for their cognate damaged bases using an extrahelical recognition mechanism. One possible contribution to specificity is the weak base pairing and inherent instability of damaged sites which may lead to increased extrahelicity of the damaged base and enhanced recognition of these sites. This model predicts that the binding affinity of the enzyme should increase as the thermodynamic stability of the lesion base pair decreases, because less work is required to extrude the base into its active site. We have tested this hypothesis with uracil DNA glycosylase (UDG) by constructing a series of DNA duplexes containing a single uracil (U) opposite a variety of bases (X) that formed from zero to three hydrogen bonds with U. Linear free energy (LFE) relationships were observed that correlated UDG binding affinity with the entropy and enthalpy of duplex melting, and the dynamic accessibility of the damaged site to chemical oxidation. These LFEs indicate that the increased conformational freedom of the damaged site brought about by enthalpic destabilization of the base pair promotes the formation of extrahelical states that enhance specific recognition by as much as 3000-fold. However, given the small stability differences between normal base pairs and U•A or U•G base pairs, relative base pair stability contributes little to the >10⁶-fold discrimination of UDG for uracil sites in cellular DNA. In contrast, the intrinsic instability of other more egregious DNA lesions may contribute significantly to the specificity of other DNA repair enzymes that bind to extrahelical bases.

The genetic information of a cell can be irreversibly altered through the chemical modification of nucleotide bases (1). To combat these mutagenic effects, organisms have evolved a two-tiered base excision repair (BER) pathway that handles a wide array of base lesions (2). In the first stage, a highly specific DNA repair glycosylase excises the damaged base from the DNA (3, 4), producing an abasic site. This intermediate is then processed by the sequential action of several repair enzymes that ultimately restore the site to its original state (2). Since damaged site specificity resides solely with the DNA glycosylase, these enzymes must possess extraordinarily high catalytic specificities (4, 5). In the absence of such specificity, undamaged bases would be randomly excised from DNA, leading to undesirable abasic sites and genetic instability (6–10).

In general, enzymatic specificity results from the extraordinary structure of enzyme active sites that disfavors formation of catalytically productive interactions with non-substrate molecules, and strongly favors such interactions with the true substrates. As an essential part of their

recognition mechanisms, all DNA glycosylases extrude their damaged bases from the DNA double helix in a process known as base flipping, thereby placing it extrahelically into their active sites where specific interactions with the damaged base can be formed (11, 12). As part of the energetic cost of base flipping, the hydrogen bonds and stacking interactions of the base pair must be disrupted. Thus, a prediction is that DNA glycosylases should bind more tightly to damaged sites with disrupted base pairing because it requires less binding energy to flip the damaged base from the destabilized site (4, 13–22). Such a thermodynamic mechanism is quite general, and would apply even for DNA glycosylases that interact with the base that opposes the damaged base (23–25).

How much does the intrinsic thermodynamic stability of the damaged base pair contribute to specific damaged site binding by DNA glycosylases? We have investigated this question using the enzyme uracil DNA glycosylase (UDG),¹ which removes uracil from U•G and U•A base pairs in duplex DNA (14, 26, 27). The approach was to measure the binding affinity of UDG for a series of DNA duplexes, in which the number of hydrogen bonds (*n*) between uracil and its opposing base (X) were systematically varied (Figure 1A).

[†] This work was supported by NIH Grant GM56834 to J.T.S.

^{*} To whom correspondence should be addressed: Department of Pharmacology and Molecular Sciences, Johns Hopkins School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205. E-mail: jstivers@jhmi.edu.

[‡] The Johns Hopkins School of Medicine.

[§] The Center for Advanced Research in Biotechnology and National Institute of Standards and Technology.

¹ Abbreviations: UDG, uracil DNA glycosylase; U^F, 2'-β-fluoro-2'-deoxyuridine; φ, abasic site; D, 2,6-diaminopurine; M, 4-methylindole; N, nebularine; DSC, differential scanning calorimetry; ψ, pseudo-dihedral angle.

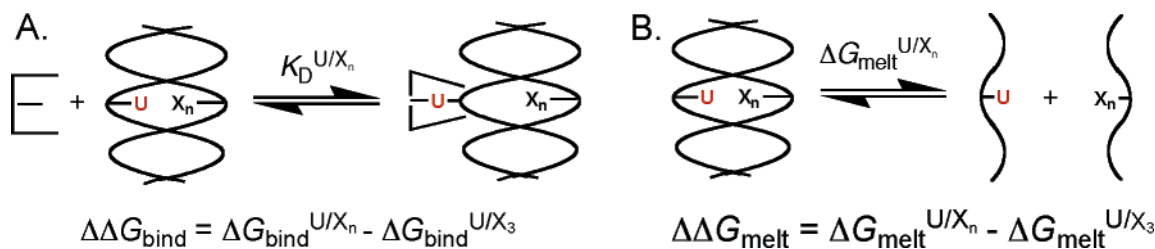


FIGURE 1: (A) Specific binding of UDG (E) to a DNA duplex containing deoxyuridine (U) opposite a purine analogue (X). In this study, the number of hydrogen bonds (n) in the $\text{U}\cdot\text{X}_n$ base pair varies from zero to three. Since binding of uracil requires breaking of the base pair hydrogen bonds, the difference in binding free energy ($\Delta\Delta G_{\text{bind}}$) between a DNA duplex with n hydrogen bonds and a reference duplex with three hydrogen bonds ($\text{U}\cdot\text{X}_3$) should in part reflect the reduced energetic cost of breaking the hydrogen bonds in the base pair. (B) Dissociation of a duplex with a $\text{U}\cdot\text{X}_n$ base pair into two single strands. The difference in the duplex melting free energy ($\Delta\Delta G_{\text{melt}}$) between a $\text{U}\cdot\text{X}_n$ duplex and the reference duplex ($\text{U}\cdot\text{X}_3$) will also reflect the energetic effects of a destabilized base pair.

Then, these binding affinities were correlated with rigorous measurements of the thermodynamic stabilities of these duplexes (Figure 1B). UDG was selected for this study because, unlike many DNA glycosylases (23–25), it does not make any direct contacts with the base opposite uracil (11). Thus, the observed changes in binding affinity can be largely attributed to the relative stability of the $\text{U}\cdot\text{X}_n$ base pairs in the free duplex DNA and not differential interactions of the enzyme with the various opposing bases. These quantitative free energy correlations provide the first direct evidence that promotion of extrahelical conformations by enthalpic destabilization of a damaged site can indeed enhance the specific binding of a DNA repair enzyme.

EXPERIMENTAL PROCEDURES²

Materials. The 2'-deoxynucleoside phosphoramidites, CPG supports, and DNA synthesis reagents were purchased from Glen Research (Sterling, VA), except for 2'- β -fluoro-2'-deoxyuridine (U^{F}), which was synthesized as described previously (5, 28). The oligonucleotides were synthesized using standard phosphoramidite chemistry on an Applied Biosystems 392 synthesizer. The oligonucleotides were purified by anion exchange HPLC (Zorbax), followed by C-18 reversed phase HPLC (Phenomenex Aqua column). Fractions containing pure oligonucleotide were concentrated to dryness *in vacuo*, redissolved in MilliQ water, and stored at -20°C . The purity of the oligonucleotides was assessed by matrix-assisted laser desorption mass spectroscopy and denaturing polyacrylamide gel electrophoresis. The concentration of each oligonucleotide was determined using its extinction coefficient at 260 nm (29). DNA duplexes were hybridized in 10 mM Tris-HCl (pH 8.0) and 25 mM NaCl as described previously (5). The purification of *Escherichia coli* UDG has been described previously (30).

K_D Measurements. The K_D values for binding of the $\text{U}^{\text{F}}\cdot\text{X}_n$ duplexes to UDG were measured essentially as described using a kinetic competitive inhibition HPLC assay under conditions where the apparent K_i is equal to the K_D value (i.e., $[\text{S}] \ll K_m$, where S exhibits rapid equilibrium binding) (31). The only modification was that the abasic product ($\text{A}\Phi\text{Ap}$) and reactant (AUAp) were separated using isocratic

conditions with 9.5% CH_3CN and 0.1 M triethylammonium acetate. Reaction mixtures (35 μL) containing 10 mM Tris-HCl (pH 8.0), 60 mM NaCl, 12.5 $\mu\text{g}/\text{mL}$ BSA, 1 μM AUAp, 0.5 nM UDG, and a variable amount of the $\text{U}^{\text{F}}\cdot\text{X}_n$ duplex were incubated at room temperature for 5 min. The K_D for each duplex was determined by fitting to eq 1

$$k_i/k_o = 1/(1 + [\text{U}^{\text{F}}\cdot\text{X}_n]/K_D) \quad (1)$$

where k_i is the inhibited rate and k_o is the rate in the absence of competitor DNA. For the tightest binding duplex ($\text{U}^{\text{F}}\cdot\text{M}$), eq 1 was modified to take into account inhibitor depletion (32).

The differences in binding free energies relative to the duplex with three hydrogen bonds ($\text{U}^{\text{F}}\cdot\text{D}$) were calculated from the measured K_D values using eq 2.

$$\Delta\Delta G_{\text{bind}} = RT \ln[K_D(\text{U}^{\text{F}}\cdot\text{X}_n)/K_D(\text{U}^{\text{F}}\cdot\text{D})] \quad (2)$$

Fluorescence Spectroscopy. To ascertain that all of the DNA duplexes attained the same bound state, tryptophan fluorescence measurements of free and DNA-bound UDG were performed. Samples (497 μL) containing 10 mM Tris-HCl (pH 8.0), 60 mM NaCl, and 300 nM UDG were incubated for 3 min at 25°C in a 10 mm quartz cuvette, and a fluorescence emission spectrum was recorded in the range of 325–425 nm on a SPEX FluoroMax-3 fluorimeter ($\lambda_{\text{ex}} = 295$ nm). $\text{U}^{\text{F}}\cdot\text{X}_n$ duplex DNA (2.5 μL) was then added to the UDG solution to give a final DNA concentration of 500 nM. The reaction mixture was magnetically stirred and incubated for 3 min at 25°C , before the fluorescence emission spectrum of the UDG·DNA complex was recorded. The tryptophan fluorescence intensities of free and DNA-bound UDG at 333 nm were measured, and the raw values were then normalized for the fraction of UDG bound to each DNA analogue before the ratio ($F_{\text{bound}}^{333}/F_{\text{free}}^{333}$) was calculated.

Differential Scanning Calorimetry. DSC measurements of duplex strand melting were taken using a VP-DSC microcalorimeter from Microcal, Inc. (Northampton, MA) essentially as described previously (33). The DNA solutions had a concentration of 20 μM in DNA duplex with 10 mM Na_2HPO_4 (pH 7.5) and 60 mM NaCl. Samples were equilibrated at 20°C for 15 min and scanned up to 95°C at a preset scan rate of $60^\circ\text{C}/\text{h}$. The transition peak areas were measured using the EXAM software program (34), and the transition peak areas were divided by the DNA duplex concentration to provide the transition enthalpies. Transition

² Certain commercial materials, instruments, and equipment are identified herein to specify the experimental procedure as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material, instruments, or equipment identified is necessarily the best available for the purpose.

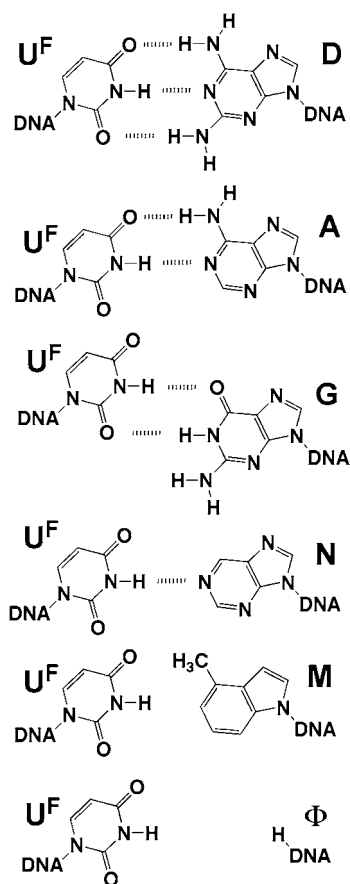


FIGURE 2: Structures of $U^F \cdot X_n$ base pairs: $U^F \cdot D$ (54, 55), $U^F \cdot G$ (56), $U^F \cdot N$ (57, 58), $U^F \cdot \Phi$ (59–61), and $U^F \cdot M$ (62).

entropies were determined from plots of C_p/T against T , by dividing the transition peak area by the DNA duplex concentration. Procedures for buffer baseline corrections and accounting for pre- and post-transition baselines have been described previously (33).

KMnO₄ Oxidation Measurements. Because of the low reactivity of U^F to oxidation (35), it was replaced with thymine (T) in the oligonucleotides used in this study. To reaction mixtures (20 μ L) containing 10 mM Tris-HCl (pH 8.0), 60 mM NaCl, and either 100 nM single-stranded 5'-[³²P]T or 5'-[³²P]T· X_n duplex was added 2.5 mM $KMnO_4$. After the sample had been incubated for 3 min at room temperature, the reaction was halted by the addition of 20 μ L of a stop solution containing 1.5 M sodium acetate, 1 M 2-mercaptoethanol, and 200 μ g/mL tRNA. The samples were processed, imaged, and quantified as described previously (32).

RESULTS

Binding of UDG to Destabilized Damaged Sites. A series of 15mer duplexes were constructed in which 2'- β -fluoro-2'-deoxyuridine (U^F), a nonhydrolyzable uracil analogue (5), was placed opposite a series of bases (X) which form zero to three hydrogen bonds with U^F (Figure 2). The affinity of UDG for each $U^F \cdot X$ duplex was measured by a competitive inhibition kinetic assay in which a 3mer substrate (AUAp) is separated from the abasic product (A Φ Ap) using reverse phase HPLC (Figure 3A) (31). Representative inhibition data are shown in Figure 3B for the duplex that contains a uracil·4-methylindole base pair ($U^F \cdot M$). The K_D values for

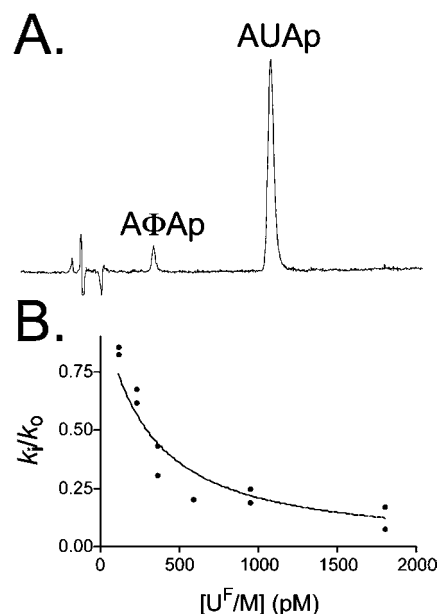


FIGURE 3: Determination of the binding affinity of specific DNA duplexes using a competitive inhibition assay. (A) UDG HPLC activity assay. The substrate (AUAp) and product A Φ Ap are indicated. (B) Inhibition of UDG by a $U \cdot M$ duplex ($K_D = 0.24 \pm 0.03$ nM).

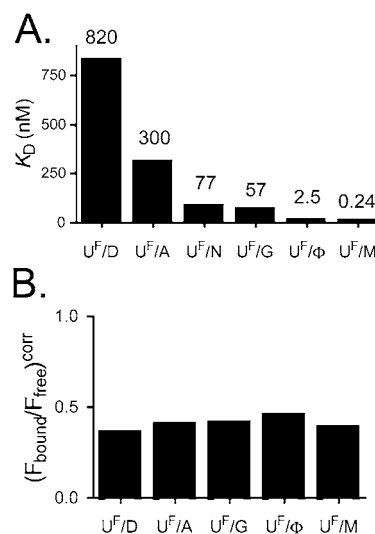


FIGURE 4: (A) Binding affinities of UDG for $U^F \cdot X_n$ DNA duplexes. The K_D values (nanomolar) are shown above the individual bars. (B) Quench of UDG tryptophan fluorescence upon binding of the $U^F \cdot X_n$ duplex. The degree of UDG quenching was normalized to reflect the quenching when UDG is saturated with DNA (see Experimental Procedures).

the six duplexes showed an incremental decrease as the number of hydrogen bonds was ablated (Figure 4A). The tightest observed binding affinity was for the $U^F \cdot M$ construct that has no hydrogen bonds ($K_D = 0.25$ nM). This affinity is 3000-fold tighter than that of the uracil·diaminopurine duplex that has three hydrogen bonds ($U^F \cdot D$, $K_D = 820$ nM). Similar tight binding was observed for the duplex with an abasic site opposite U^F ($U^F \cdot \Phi$, $K_D = 2.4$ nM), providing additional evidence that complete ablation of hydrogen bonding leads to a significantly increased binding affinity.³

³ The observed K_D for the $U^F \cdot \Phi$ duplex (2.5 nM) predominantly reflects binding of UDG to U^F , and not to Φ , because the K_D for abasic DNA is much weaker (70 nM for T· Φ and 100 nM for $\Phi \cdot A$).

Table 1: Thermodynamic Parameters for UDG Binding, DNA Melting, and Permanganate Accessibility of Destabilized Base Pairs

duplex	K_D (nM)	ΔG_{bind} (kcal/mol)	ΔH_{melt} (kcal/mol)	ΔS_{melt} (cal mol ⁻¹ K ⁻¹)	ΔG_{melt}^a (kcal/mol)	log S^b
U ^F •D	820 ± 90	-8.3 ± 0.1	58.9 ± 1.8	0.16 ± 0.01	10.9 ± 2.3	-1.4 ± 0.14
U ^F •A	300 ± 50	-8.9 ± 0.1	39.6 ± 4.0	0.11 ± 0.01	7.1 ± 5.2	-1.0 ± 0.06
U ^F •N	77 ± 11	-9.7 ± 0.1	40.5 ± 2.0	0.10 ± 0.01	9.2 ± 2.5	ND ^c
U ^F •G	57 ± 6	-9.9 ± 0.1	34.0 ± 2.4	0.10 ± 0.01	5.6 ± 3.2	-0.80 ± 0.03
U ^F •Φ ³	2.5 ± 0.4	-11.7 ± 0.1	4.3 ± 0.1	0.01 ± 0.01	0.3 ± 0.3	-0.22 ± 0.01
U ^F •M	0.24 ± 0.03	-13.1 ± 0.1	ND ^c	ND ^c	ND ^c	-0.36 ± 0.01

^a Calculated at 298 K. ^b S is the relative sensitivity of a T•X base pair to oxidation by KMnO₄ (see the legend of Figure 6). ^c Not determined.

These results indicate that removal of three hydrogen bonds can enhance specific recognition by up to 4.8 kcal/mol.

In large part, these binding effects reflect the thermodynamic properties of the free damaged site because these discrete base pair perturbations are not expected to affect interactions between UDG and the DNA.⁴ This conclusion is supported by inspection of the crystal structures of UDG complexed with substrate analogues, which show that UDG does not make any interactions with the base that opposes the damage site, or with the undamaged strand (36). Thus, recognition solely involves the extrahelical deoxyuridine and not other specific features of the base pair or duplex. To further establish that all of the U^F•X duplexes used here attain the same bound state, and that the observed effects on binding largely arise from the properties of the free DNA, we measured the tryptophan fluorescence quenching upon binding of each duplex (Figure 4B). Previous work has shown that the quenching of UDG tryptophan fluorescence upon specific DNA binding is a sensitive measure of an induced fit conformational change in UDG that is required to achieve the final productive conformation with a flipped-out uracil (5). Within the errors of these measurements, all of the duplexes produced the same magnitude of fluorescence quenching, indicating that the same bound conformations were attained for all.

Thermodynamic Stabilities of U^F•X Duplexes. We then determined the energetic effects of this series of site-specific base pair disruptions on the thermodynamic parameters for DNA melting using differential scanning calorimetry (DSC, Figure 5). With DSC, one can measure the enthalpy (ΔH_{melt}) and entropy (ΔS_{melt}) of the melting transition directly, and unlike optical methods, it is insensitive to the mechanism of duplex melting (37). The free energy of duplex dissociation (ΔG_{melt}) at any temperature can then be simply calculated from ΔH_{melt} and ΔS_{melt} , using the relationship $\Delta G_{\text{melt}} = \Delta H_{\text{melt}} - T\Delta S_{\text{melt}}$, because the heat capacity of the duplex and that of the single strands are equal ($\Delta C_p \approx 0$) (38, 39). As expected, U^F•X duplexes exhibited decreasing transition enthalpies in the ΔH_{melt} range of 58.9–4.3 kcal/mol as the number of hydrogen bonds was decreased. The complete thermodynamic parameters for duplex melting are reported in Table 1.

Correlation of UDG Binding Affinity with Damaged Base Pair Stability. To quantitatively evaluate the impact of damaged base pair disruption on UDG affinity, the differ-

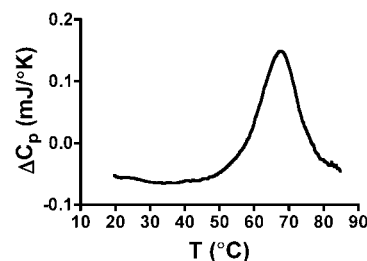


FIGURE 5: Differential scanning calorimetry (DSC) of the U^F•D duplex. A 20 μM DNA duplex solution in 10 mM sodium phosphate (pH 7.5) and 60 mM NaCl was placed inside of the DSC sample cell, and the change in the heat capacity of the solution (ΔC_p) was monitored as it was warmed from 20 to 95 °C at a rate of 60 °C/h. The thermodynamic parameters of duplex dissociation were extracted as described in Experimental Procedures.

ences in binding free energies ($\Delta\Delta G_{\text{bind}}$, Figure 1) were plotted against the changes in transition enthalpies ($\Delta\Delta H_{\text{melt}}$) and entropies ($-T\Delta\Delta S_{\text{melt}}$) (Figure 6A,B). In this analysis, the difference energies are relative to the duplex with three hydrogen bonds (U^F•D), and the value of $-T\Delta\Delta S_{\text{melt}}$ was calculated at 298 K. $\Delta\Delta G_{\text{bind}}$ was found to increase linearly as the transition enthalpy and entropy decrease: $\Delta\Delta H_{\text{melt}}$ (slope = 0.064 ± 0.01 , $r^2 = 0.937$) and $-T\Delta\Delta S_{\text{melt}}$ (slope = 0.080 ± 0.011 , $r^2 = 0.949$). As required from these linear correlations, a strong correlation with $\Delta\Delta G_{\text{melt}}$ (slope = 0.3 ± 0.1 , $r^2 = 0.839$) was also observed (data not shown). The implications of these correlations are discussed below.

Correlation of UDG Binding Affinity with Damaged Base Accessibility. The above thermodynamic correlations suggest that these destabilized duplexes might exhibit an increase in the number of dynamic fluctuations that promote extrahelical states of U^F at temperatures well below the duplex melting temperature. Since the dynamic accessibility of the damaged base is another factor that could enhance its recognition by DNA glycosylases, it was of interest to measure the relative accessibility of each destabilized base pair, and correlate this parameter with UDG binding affinity.

To explore this question, a potassium permanganate (KMnO₄) sensitivity assay was employed (32, 35, 40). Since sites of pyrimidine oxidation are susceptible to strand cleavage under basic conditions, they can be detected as fragments using polyacrylamide gel electrophoresis (Figure 7A) (41). In these studies, U^F was first replaced with thymine (T) because of its poor oxidation reactivity arising from its electron deficient 5,6-double bond (data not shown). This is a very conservative change, as a T•X base pair will have a base pairing strength and geometry nearly identical to those of a U^F•X base pair (42, 43). As shown in Figure 7A, decreasing the base pair strength leads to an increase in the sensitivity of the T•X base pair to permanganate oxidation, and also the invariant T•A base pair three nucleotides away,

⁴ The observation of linear free energy relationships between the thermodynamic parameters of the free duplex DNA and the overall free energy of DNA binding is not negated if different energetic interactions exist between the enzyme and each bound duplex. However, the observed slopes will reflect the relative effects of perturbing the duplex in the free and bound state.

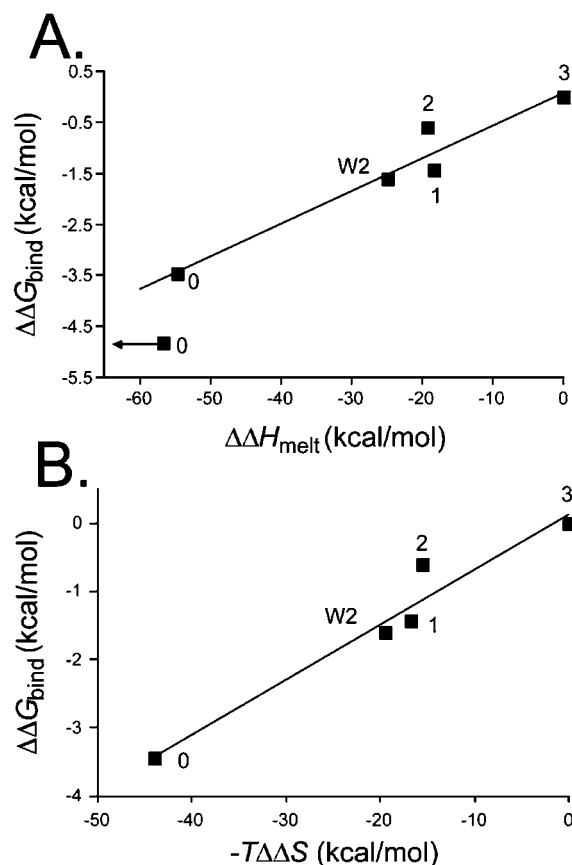


FIGURE 6: Correlations between UDG binding affinity for each $U^F \cdot X_n$ duplex ($\Delta\Delta G_{\text{bind}}$) and the thermodynamic parameters for duplex melting. The numbers of hydrogen bonds in the $U^F \cdot X_n$ base pair are indicated. (A) $\Delta\Delta G_{\text{bind}}$ vs $\Delta\Delta H_{\text{melt}}$ (slope = 0.064, $r^2 = 0.937$). Only an upper limit for the enthalpy of melting for the $U^F \cdot M$ duplex was obtained (arrow). (B) $\Delta\Delta G_{\text{bind}}$ vs $-T\Delta\Delta S_{\text{melt}}$ (slope = 0.080, $r^2 = 0.949$).

indicating that even very conservative changes to the $T \cdot X$ base pair can influence the dynamics of neighboring base pairs in the duplex (44). A plot of $\log K_D$ against $\log(\text{relative KMnO}_4 \text{ sensitivity})$ shows a linear correlation (Figure 7B), establishing that increasing base accessibility at temperatures well below the T_m value has a strong positive effect on binding affinity. As an important control, the amount of oxidized product was found to increase linearly with respect to time and concentration of KMnO_4 (data not shown). Therefore, the differences in the sensitivity of these duplexes to oxidation directly reflect the unfavorable dynamic pre-equilibrium for exposure of the thymidine base prior to reaction with KMnO_4 .

DISCUSSION

Thermodynamic Framework for Active and Passive Base Flipping. While it is clear from structural and spectroscopic studies that DNA glycosylases bind their cognate damaged base in an extrahelical conformation (3, 11), the pathway by which the damaged base is flipped out of the DNA duplex and placed inside of the enzyme active site remains poorly defined. In one model, DNA glycosylases *passively* capture damaged bases that are transiently extrahelical (Figure 8) (14, 45). According to this view, the DNA glycosylase does not lower the activation energy or equilibrium for damaged base flipping, but instead relies on the increased extrahelical propensity of damaged bases to enhance bimolecular en-

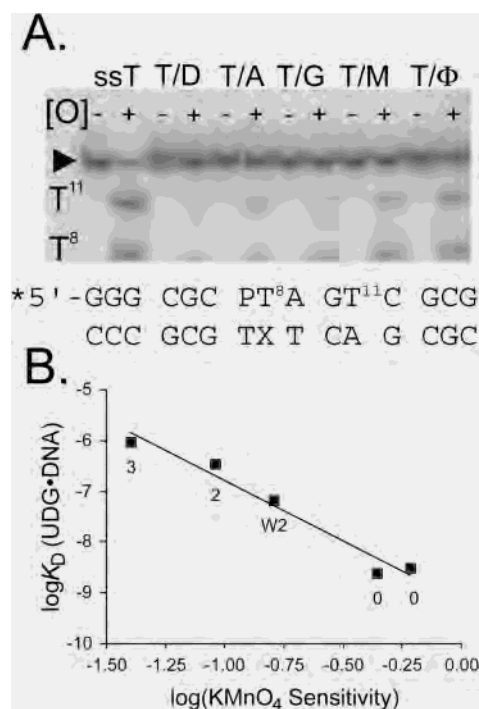


FIGURE 7: (A) Sensitivity of thymine in single-stranded and duplex DNA to oxidation by potassium permanganate. Samples of $5' \cdot ^{32}\text{P}$ -labeled single-stranded or $T \cdot X_n$ duplex DNAs were reacted with 2.5 mM KMnO_4 ([O]) for 3 min. After the oxidized strands had been cleaved with piperidine, the reaction mixtures were run a 19% denaturing polyacrylamide gel, and the radioactivity of each band was quantified with a phosphorimager. The position on the gel of the full-length DNA is marked with a black wedge. (B) Correlation between UDG binding affinity and KMnO_4 sensitivity (slope = -2.37 ± 0.270 , $r^2 = 0.963$). The relative sensitivity of a thymine in a $T \cdot X_n$ base pair is defined as $[(I^{T8} - I^{\text{bgd}})/(I^{\text{total}} - I^{\text{bgd}})_{\text{duplex}} \times 100]/[(I^{T8} - I^{\text{bgd}})/(I^{\text{total}} - I^{\text{bgd}})_{\text{single-stranded}} \times 100]$, where I^{T8} is the intensity of the band corresponding to oxidation of the T in the $T \cdot X_n$ base pair, I^{bgd} is the background correction, and I^{total} is the sum of all of the intensities of the bands in a given lane. The numbers of hydrogen bonds in the $T \cdot X_n$ duplex are indicated.

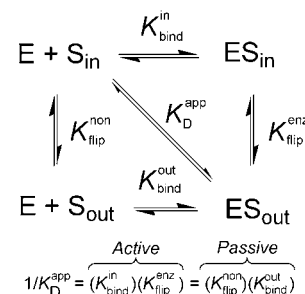


FIGURE 8: Thermodynamic model showing the energetic equivalence of the passive and active base flipping mechanisms (see the text). The equivalence of the pathways is illustrated by the thermodynamic box, which requires that $1/K_D^{\text{app}} = K_{\text{in}}^{\text{bind}} K_{\text{enz}}^{\text{enz}} = K_{\text{flip}}^{\text{non}} K_{\text{out}}^{\text{bind}}$.

counter. In the alternative view, DNA glycosylases *actively* flip out their cognate lesions by destabilizing the damaged site in an initial encounter complex. Active base flipping may occur by stabilization of high-energy intermediate conformations on the base flipping pathway, or by the use of mechanical forces to propel the base from the duplex (Figure 8) (5, 46). It is important to point out that passive and active flipping pathways cannot be distinguished by *thermodynamic measurements alone*. This conclusion is required because $K_{\text{in}}^{\text{bind}} K_{\text{enz}}^{\text{enz}} = K_{\text{flip}}^{\text{non}} K_{\text{out}}^{\text{bind}}$, as shown in

Figure 8. Only an assessment of the *kinetic competence* of each pathway can reveal whether passive or active base flipping is the major route taken for a given system.

Specific Recognition and Conformational Freedom of Damaged Sites. We have found that the affinity of UDG for a specific site is linearly dependent on the enthalpy of duplex dissociation (Figure 6A). The linear relationship between UDG affinity and ΔH_{melt} confirms the notion that weakened base pairing decreases the enthalpy of duplex melting and, consequently, increases UDG binding affinity (Figure 6A). However, this finding alone provides little insight into the physical mechanism by which decreases in base pair enthalpy lead to enhanced binding. The small slope of the correlation suggests that the higher affinity of UDG for destabilized damaged sites does not simply arise from the reduced enthalpic cost of breaking a destabilized base pair during base flipping, and that other energetic influences must be at work.⁴ It should be stated that the experiments presented here are explicitly designed to probe the enthalpic and entropic contributions of the damaged site alone to specific recognition. The total enthalpy and entropy of binding (including the enzyme, DNA, and solvent) are not evaluated in any of the current experiments, and in fact, these measurements are not required for the conclusions presented below.

The plot of $\Delta\Delta G_{\text{bind}}$ versus $-T\Delta\Delta S_{\text{melt}}$ reveals that there is an equally significant correlation between changes in duplex entropy and binding affinity (Figure 6B). The parameter $-T\Delta\Delta S_{\text{melt}}$ likely reflects the increased conformational flexibility of the destabilized base pairs in the duplex DNA, because the entropy differences of the dissociated single-stranded DNAs in the melting experiments should be similar, given the conservative changes in these substrates. If we assume this physical interpretation for the entropy changes between these DNA constructs, the correlation suggests that increased flexibility of the base pair produces conformational states that are productive for UDG binding. The conclusion that extrahelical conformational states are produced is supported by the correlation between UDG binding affinity and KMnO_4 sensitivity (Figure 7B), which reflects the dynamic equilibrium of the base between an inaccessible and permanganate accessible state (i.e., an extrahelical exposed conformation). A reasonable interpretation of these combined findings is that enthalpic destabilization of the base pair allows increased conformational flexibility, producing extrahelical conformers, some of which favor enzyme binding.

The effects of base pair enthalpy and entropy on the extrahelical conformational distributions that may promote base flipping are depicted in Figure 9A–C. In these panels, the probability of an extrahelical conformation is plotted against the backbone pseudodihedral angle (ψ) of the deoxyuridine nucleotide, defined as indicated in Figure 9. The angle ψ has been previously used in computational studies to describe the pathway for base flipping (47), and is used here because of its simple representation of the base flipping trajectory, although none of the arguments depend on this formalism. Using this nomenclature, a ψ of 10° reflects the fully base paired state and a ψ of 180° reflects the fully extrahelical state. In free DNA (Figure 9A), the U nucleotide in the stable base pair with three hydrogen bonds (U·D) should be tightly centered around an average conformation with a ψ of 10° (red curve), while the unstable

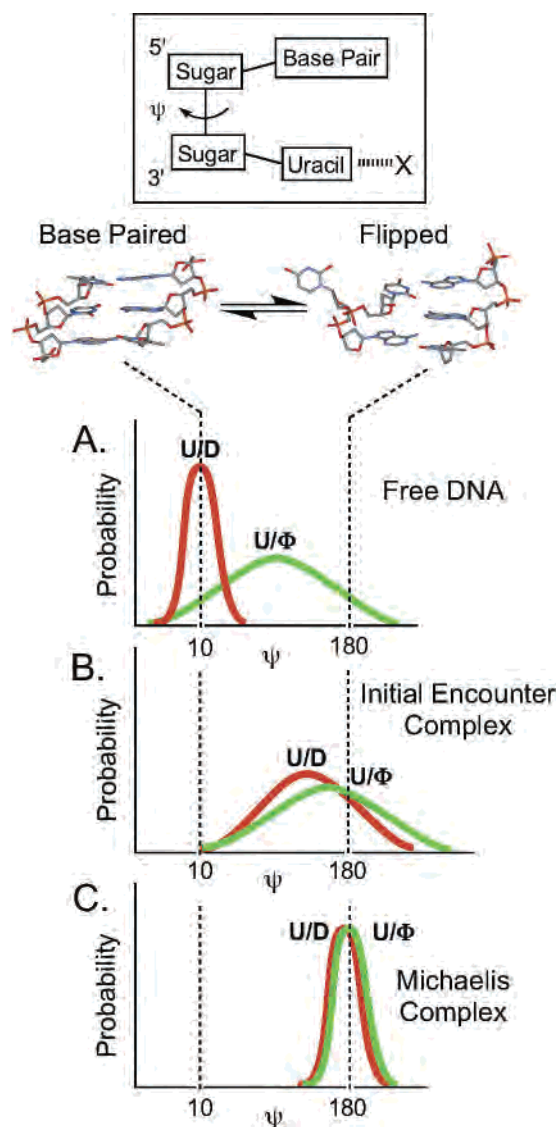


FIGURE 9: Enthalpic destabilization of the $\text{U}\cdot\text{X}_n$ base pair leads to extrahelical conformations that promote binding. The probability of an extrahelical conformer as a function of the pseudodihedral angle ψ , which is a measure of the progress along the base flipping reaction coordinate, is indicated (see the text and ref 47). Using this nomenclature, a fully stacked base pair has a ψ value of $\sim 10^\circ$, whereas the fully extrahelical conformation has a ψ value of 180° . (A) Enthalpic destabilization of the base pair in the free DNA leads to extrahelical conformations that promote *passive* base flipping. (B) Base pair destabilization can also affect an active mechanism in which UDG forms an initial encounter complex with the DNA in which the base is not yet fully extrahelical (5, 46, 48). In this case, enzyme binding energy is used to destabilize the $\text{U}\cdot\text{D}$ base pair, allowing it to achieve extrahelical states already available to the $\text{U}\cdot\Phi$ duplex due to its intrinsic instability. (C) A hypothetical distribution of extrahelical conformers for the final UDG·DNA complex. The enzyme has fully stabilized the flipped-out uracil, and the distribution of conformational states for the $\text{U}\cdot\text{D}$ and $\text{U}\cdot\Phi$ duplexes is narrowly focused around a ψ of 180° (36, 47).

construct with no hydrogen bonds ($\text{U}\cdot\Phi$) should be broadly centered around a ψ value further along the flipping reaction coordinate (green curve). Thus, enthalpic destabilization of the base pair in a passive mechanism leads to conformations that facilitate binding of the enzyme (see Figure 8, counter-clockwise pathway).

For the alternative active base flipping mechanism (Figure 9B), binding of UDG in an initial encounter complex can

alter the DNA structure such that the more stable U•D base pair is also destabilized (Figure 9B, red curve) (5, 46, 48). In this initial destabilized complex, the partially extrahelical uracil may assume a similar average conformation, and broad distribution, for both the U•D and U•Φ constructs. However, flipping of U in the U•D construct requires a greater amount of binding energy to overcome the enthalpic barrier to base flipping, resulting in weaker binding of the U•D duplex compared to that of the U•Φ duplex. Finally, in the Michaelis complex, in which the base is fully extrahelical (Figure 9C), both the U•D and U•Φ duplexes assume the same average conformation and small conformational distribution which are enforced by the strong interactions between the enzyme and uracil (36, 46). Although this mechanism implies a reduction in the conformational flexibility of destabilized base pairs such as the U•Φ duplex upon formation of the Michaelis complex, this expected unfavorable entropic contribution to the overall free energy of binding may be paid for by the even larger enthalpic benefit of base pair destabilization for such conformationally flexible substrates (Table 1).

CONCLUSIONS

We have shown that the maximum gain in specific recognition by UDG arising from destabilization of the damaged base pair is 4.8 kcal/mol, and that this energetic effect likely arises from increasing the population of extrahelical states that promote binding. Since the catalytic specificity of UDG for uracil as opposed to other normal bases has been estimated to be at least 8.3 kcal/mol *in vitro* (5), the energetic contribution of destabilized damaged sites to specific ground state binding can be significant, at least in this model system. However, the *in vivo* substrates of UDG consist of U•A or U•G base pairs, which are not significantly destabilized compared to other normal base pairs. Thus, spontaneous base flipping at damaged sites is not a viable mechanism for accounting for the specificity of UDG *in vivo*. The remaining specificity of UDG must be attributed to strong transition state interactions that can be induced only by actively flipping the uracil base into the active site (5, 49). Although unimportant for UDG, damaged site instability could contribute significantly to specific recognition by repair enzymes that act on intrinsically unstable base pairs such as *O*⁶-methylguanine (45), hypoxanthine (50, 51), *N*¹-methyladenine, and *N*³-methylcytosine (52, 53).

ACKNOWLEDGMENT

We thank Prof. Paul Miller at the Johns Hopkins School of Public Health for the generous use of his solid-phase DNA synthesizer.

REFERENCES

- Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature* 362, 709–715.
- Seeberg, E., Eide, L., and Bjoras, M. (1995) The base excision repair pathway, *Trends Biochem. Sci.* 20, 391–397.
- Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P., and Tainer, J. A. (1999) DNA repair mechanisms for the recognition and removal of damaged DNA bases, *Annu. Rev. Biophys. Biomol. Struct.* 28, 101–128.
- Stivers, J. T., and Jiang, Y. L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases, *Chem. Rev.* 103, 2729–2759.
- Stivers, J. T., Pankiewicz, K. W., and Watanabe, K. A. (1999) Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase, *Biochemistry* 38, 952–963.
- Lindahl, T., and Andersson, A. (1972) Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid, *Biochemistry* 11, 3618–3623.
- Cuniasse, P., Fazakerley, G. V., Guschlbauer, W., Kaplan, B. E., and Sowers, L. C. (1990) The abasic site as a challenge to DNA polymerase. A nuclear magnetic resonance study of G, C and T opposite a model abasic site, *J. Mol. Biol.* 213, 303–314.
- Lindahl, T. (1990) Repair of intrinsic DNA lesions, *Mutat. Res.* 238, 305–311.
- Kavli, B., Slupphaug, G., Mol, C. D., Arvai, A. S., Peterson, S. B., Tainer, J. A., and Krokan, H. E. (1996) Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase, *EMBO J.* 15, 3442–3447.
- Kwon, K., Jiang, Y., and Stivers, J. (2003) Rational Engineering of a DNA Glycosylase Specific for Unnatural Cytosine:Pyrene Base Pairs, *Chem. Biol.* 10, 1–9.
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature* 384, 87–92.
- Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix, *Cell* 76, 357–369.
- Verdine, G. L., and Bruner, S. D. (1997) How do DNA repair proteins locate damaged bases in the genome? *Chem. Biol.* 4, 329–334.
- Pearl, L. H. (2000) Structure and function in the uracil-DNA glycosylase superfamily, *Mutat. Res.* 460, 165–181.
- Vallur, A. C., Feller, J. A., Abner, C. W., Tran, R. K., and Bloom, L. B. (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase, *J. Biol. Chem.* 277, 31673–31678.
- Osman, R., Fuxreiter, M., and Luo, N. (2000) Specificity of damage recognition and catalysis of DNA repair, *Comput. Chem.* 24, 331–339.
- Fuxreiter, M., Luo, N., Jedlovsky, P., Simon, I., and Osman, R. (2002) Role of base flipping in specific recognition of damaged DNA by repair enzymes, *J. Mol. Biol.* 323, 823–834.
- Liu, P., Burdzy, A., and Sowers, L. C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG, and TDG, *Chem. Res. Toxicol.* 15, 1001–1009.
- Valinluck, V., Liu, P., Burdzy, A., Ryu, J., and Sowers, L. C. (2002) Influence of Local Duplex Stability and N(6)-Methyladenine on Uracil Recognition by Mismatch-Specific Uracil-DNA Glycosylase (Mug), *Chem. Res. Toxicol.* 15, 1595–1601.
- Biswas, T., Clos, L. J., II, SantaLucia, J., Jr., Mitra, S., and Roy, R. (2002) Binding of specific DNA base-pair mismatches by *N*-methylpurine-DNA glycosylase and its implication in initial damage recognition, *J. Mol. Biol.* 320, 503–513.
- Panayotou, G., Brown, T., Barlow, T., Pearl, L. H., and Savva, R. (1998) Direct measurement of the substrate preference of uracil-DNA glycosylase, *J. Biol. Chem.* 273, 45–50.
- Chepanoske, C. L., Langelier, C. R., Chmiel, N. H., and David, S. S. (2000) Recognition of the nonpolar base 4-methylindole in DNA by the DNA repair adenine glycosylase MutY, *Org. Lett.* 2, 1341–1344.
- Fromme, J. C., and Verdine, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, *Nat. Struct. Biol.* 9, 544–552.
- Bernards, A. S., Miller, J. K., Bao, K. K., and Wong, I. (2002) Flipping duplex DNA inside-out: A double base-flipping reaction mechanism by *Escherichia coli* MutY adenine glycosylase, *J. Biol. Chem.* 277, 20960–20964.
- Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., and Pearl, L. H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions, *Cell* 92, 117–129.
- Lindahl, T. (1974) An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues, *Proc. Natl. Acad. Sci. U.S.A.* 71, 3649–3653.

27. Stivers, J. T., and Drohat, A. C. (2001) Uracil DNA glycosylase: insights from a master catalyst, *Arch. Biochem. Biophys.* **396**, 1–9.
28. Watanabe, K. A., Reichman, U., Hirota, K., Lopez, C., and Fox, J. J. (1979) Nucleosides. 110. Synthesis and antihelical virus activity of some 2'-fluoro-2'-deoxyarabinofuranosylpyrimidine nucleosides, *J. Med. Chem.* **22**, 21–24.
29. Fasman, G. D. (1975) *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed., Vol. 1, CRC Press, Boca Raton, FL.
30. Drohat, A. C., Jagadeesh, J., Ferguson, E., and Stivers, J. T. (1999) The role of electrophilic and base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase, *Biochemistry* **38**, 11866–11875.
31. Jiang, Y. L., and Stivers, J. T. (2001) Reconstructing the substrate for uracil DNA glycosylase: tracking the transmission of binding energy in catalysis, *Biochemistry* **40**, 7710–7719.
32. Jiang, Y. L., Kwon, K., and Stivers, J. T. (2001) Turning on uracil-DNA glycosylase using a pyrene nucleotide switch, *J. Biol. Chem.* **276**, 42347–42354.
33. Chakrabarti, M. C., and Schwarz, F. P. (1999) Thermal stability of PNA/DNA and DNA/DNA duplexes by differential scanning calorimetry, *Nucleic Acids Res.* **27**, 4801–4806.
34. Kirchoff, W. H. (1993) Exam: A Two-State Thermodynamic Analysis Program, NIST Technical Note 1401, pp 1–103, National Institute of Standards and Technology.
35. Jones, A. S., and Walker, R. T. (1963) The permanganate oxidation of nucleosides, *J. Chem. Soc.*, 3554–3557.
36. Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M., and Tainer, J. A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5083–5088.
37. Breslauer, K. J. (1995) Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions, *Methods Enzymol.* **259**, 221–242.
38. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex, *Biochemistry* **34**, 16148–16160.
39. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3746–3750.
40. Hayatsu, H., and Ukita, T. (1967) The selective degradation of pyrimidines in nucleic acids by permanganate oxidation, *Biochem. Biophys. Res. Commun.* **29**, 556–561.
41. Rubin, C. M., and Schmid, C. W. (1980) Pyrimidine-specific chemical reactions useful for DNA sequencing, *Nucleic Acids Res.* **8**, 4613–4619.
42. Delort, A. M., Neumann, J. M., Molko, D., Herve, M., Teoule, R., and Tran Dinh, S. (1985) Influence of uracil defect on DNA structure: ¹H NMR investigation at 500 MHz, *Nucleic Acids Res.* **13**, 3343–3355.
43. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
44. Moe, J. G., and Russu, I. M. (1992) Kinetics and energetics of base-pair opening in 5'-d(CGCGAATTCGCG)-3' and a substituted dodecamer containing G-T mismatches, *Biochemistry* **31**, 8421–8428.
45. Duguid, E. M., Mishina, Y., and He, C. (2003) How Do DNA Repair Proteins Locate Potential Base Lesions? A Chemical Crosslinking Method to Investigate O(6)-Alkylguanine-DNA Alkyltransferases, *Chem. Biol.* **10**, 827–835.
46. Huang, N., Banavali, N. K., and MacKerell, A. D., Jr. (2003) Protein-facilitated base flipping in DNA by cytosine-5-methyltransferase, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 68–73.
47. Banavali, N. K., and MacKerell, A. D., Jr. (2002) Free energy and structural pathways of base flipping in a DNA GCGC containing sequence, *J. Mol. Biol.* **319**, 141–160.
48. Jiang, Y. L., and Stivers, J. T. (2002) Mutational analysis of the base flipping mechanism of uracil DNA glycosylase, *Biochemistry* **41**, 11236–11247.
49. Jiang, Y. L., Drohat, A. C., Ichikawa, Y., and Stivers, J. T. (2002) Probing the Limits of Electrostatic Catalysis by Uracil DNA Glycosylase Using Transition-State Mimicry and Mutagenesis, *J. Biol. Chem.* **277**, 15385–15392.
50. Case-Green, S. C., and Southern, E. M. (1994) Studies on the base pairing properties of deoxyinosine by solid phase hybridization to oligonucleotides, *Nucleic Acids Res.* **22**, 131–136.
51. Wyatt, M. D., and Samson, L. D. (2000) Influence of DNA structure on hypoxanthine and 1,N(6)-ethenoadenine removal by murine 3-methyladenine DNA glycosylase, *Carcinogenesis* **21**, 901–908.
52. Treweek, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage, *Nature* **419**, 174–178.
53. Farnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*, *Nature* **419**, 178–182.
54. Cheong, C., Tinoco, I., Jr., and Chollet, A. (1988) Thermodynamic studies of base pairing involving 2,6-diaminopurine, *Nucleic Acids Res.* **16**, 5115–5122.
55. Howard, F. B., Frazier, J., and Miles, H. T. (1966) A new polynucleotide complex stabilized by 3 interbase hydrogen bonds, poly-2-aminoadenylic acid + polyuridylic acid, *J. Biol. Chem.* **241**, 4293–4295.
56. Carbonnaux, C., Fazakerley, G. V., and Sowers, L. C. (1990) An NMR structural study of deaminated base pairs in DNA, *Nucleic Acids Res.* **18**, 4075–4081.
57. Ikuta, S., Eritja, R., Kaplan, B. E., and Itakura, K. (1987) NMR studies of the stable mismatch purine-thymine in the self-complementary d(CGPaAATTCG) duplex in solution, *Biochemistry* **26**, 5646–5650.
58. Clore, G. M., Oschkinat, H., McLaughlin, L. W., Benseler, F., Happ, C. S., Happ, E., and Gronenborn, A. M. (1988) Refinement of the solution structure of the DNA dodecamer 5'-d(CGCGPATTCGCG)₂ containing a stable purine-thymine base pair: combined use of nuclear magnetic resonance and restrained molecular dynamics, *Biochemistry* **27**, 4185–4197.
59. Cuniasse, P., Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., Cognet, J. A., LeBret, M., Guschlbauer, W., and Fazakerley, G. V. (1987) An abasic site in DNA. Solution conformation determined by proton NMR and molecular mechanics calculations, *Nucleic Acids Res.* **15**, 8003–8022.
60. Goljer, I., Kumar, S., and Bolton, P. H. (1995) Refined solution structure of a DNA heteroduplex containing an aldehydic abasic site, *J. Biol. Chem.* **270**, 22980–22987.
61. Coppel, Y., Berthet, N., Coulombeau, C., Coulombeau, C., Garcia, J., and Lhomme, J. (1997) Solution conformation of an abasic DNA undecamer duplex d(CGACXCACGC) × d(GCGTGTGTGCG): the unpaired thymine stacks inside the helix, *Biochemistry* **36**, 4817–4830.
62. Moran, S., Ren, R. X., Sheils, C. J., Rumney, S. t., and Kool, E. T. (1996) Non-hydrogen bonding “terminator” nucleosides increase the 3'-end homogeneity of enzymatic RNA and DNA synthesis, *Nucleic Acids Res.* **24**, 2044–2052.

BI036303Y

Articles

The Origins of High-Affinity Enzyme Binding to an Extrahelical DNA Base[†]

Daniel J. Krosky,[‡] Fenhong Song,[§] and James T. Stivers^{*‡}

Department of Pharmacology and Molecular Sciences, The Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205-2185, and Center for Advanced Research in Biotechnology of the University of Maryland Biotechnology Institutes and the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received January 14, 2005; Revised Manuscript Received March 3, 2005

ABSTRACT: Base flipping is a highly conserved strategy used by enzymes to gain catalytic access to DNA bases that would otherwise be sequestered in the duplex structure. A classic example is the DNA repair enzyme uracil DNA glycosylase (UDG) which recognizes and excises unwanted uracil bases from DNA using a flipping mechanism. Previous work has suggested that enzymatic base flipping begins with dynamic breathing motions of the enzyme-bound DNA substrate, and then, only very late during the reaction trajectory do strong specific interactions with the extrahelical uracil occur. Here we report that UDG kinetically and thermodynamically prefers substrate sites where the uracil is paired with an unnatural adenine analogue that lacks any Watson–Crick hydrogen-bonding groups. The magnitude of the preference is a striking 43000-fold as compared to an adenine analogue that forms three H-bonds. Transient kinetic and fluorescence measurements suggest that preferential recognition of uracil in the context of a series of incrementally destabilized base pairs arises from two distinct effects: weak or absent hydrogen bonding, which thermodynamically assists extrusion, and, most importantly, increased flexibility of the site which facilitates DNA bending during base flipping. A coupled, stepwise reaction coordinate is implicated in which DNA bending precedes base pair rupture and flipping.

Enzymes that modify or cleave nucleobases in DNA, such as DNA methyltransferases and DNA glycosylases, are confronted with a formidable chemical problem: gaining access to substrate bases that are sequestered inside the DNA double helix (1–3). A conserved enzymatic solution to this problem is base flipping, where the target base and sugar

are extruded from the DNA duplex into the enzyme active site (4).

Structural and mechanistic studies indicate that base flipping is a multistep process involving two coupled reaction coordinates (Figure 1A) (5–12), which are depicted using a free energy contour plot in Figure 1B. The first coordinate involves $\sim 180^\circ$ rotation of the entire target nucleotide from the duplex stack (vertical axis, Figure 1B), while the second coordinate involves enzyme-induced DNA bending (horizontal axis, Figure 1B). The progress along each reaction coordinate may be perfectly synchronized (diagonal dashed line, Figure 1B), or alternatively, one process may proceed ahead of the other (curved trajectory, Figure 1B). For instance, if DNA bending proceeds ahead of nucleotide

[†] This work was supported by National Institutes of Health Grant GM56834-10. D.J.K. was supported by the DOD Breast Cancer Research Program (DAMD17-03-1-1251).

^{*} To whom correspondence should be addressed. Tel: 410-502-2758. Fax: 410-955-3023. E-mail: jstivers@jhmi.edu.

[‡] The Johns Hopkins School of Medicine.

[§] Center for Advanced Research in Biotechnology of the University of Maryland Biotechnology Institutes and the National Institute of Standards and Technology.

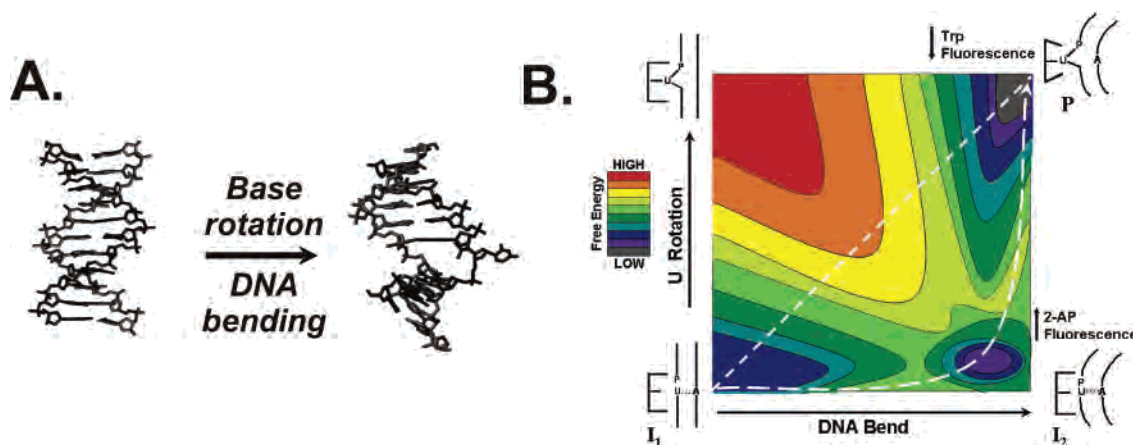


FIGURE 1: Uracil flipping has two coupled reaction coordinates involving base rotation and DNA bending. (A) Upon binding to free B DNA (left) UDG rotates the deoxyuridine nucleotide from the DNA base stack by 180° and bends the DNA by about 40° . The DNA structure on the right was extracted from the complex of UDG with substrate analogue DNA (PDB code 1EMH) (8). (B) A two-dimensional free energy contour map depicting the two coupled reaction coordinates of base rotation (vertical coordinate) and DNA bending (horizontal coordinate). The diagram shows the two enzyme-bound intermediates (I_1 , I_2) that have been previously detected using rapid kinetic methods (5, 6) and the final product of the flipping and bending reaction (P). In principle, progress along both reaction coordinates may be synchronized (diagonal line), or one process may lag behind the other. The contour map depicts the reaction trajectory (curved dashed line) where bending precedes flipping, and a low energy (blue) bent intermediate is formed before significant progress along the base rotation coordinate occurs (I_2). Thus, flipping uracil from unbent DNA is a high-energy (red) improbable process (upper left corner). Formation of I_2 can be followed using the increased fluorescence of the 2-aminopurine probe which is adjacent to the uracil and is very sensitive to base stacking (see Figure 2) (5, 6). Formation of P is discretely monitored by following the decrease in tryptophan fluorescence of UDG that accompanies base flipping (5, 6).

rotation, a bent intermediate may form (lower right corner, Figure 1B) before the hydrogen bonds to the target base are broken and the base is expelled into the enzyme active site (upper right corner, Figure 1B). A coupled reaction coordinate for base flipping suggests that if discrete alterations in the structural or dynamic features of the DNA substrate are made, then progress along one or both coordinates could be perturbed in a systematic way. This would allow a linear free energy perturbation (LFEP)¹ analysis analogous to that employed in simple chemical reactions involving coupled processes such as bond formation to a nucleophile and bond breakage to a leaving group (13).

A LFEP approach that we are continuing to explore for the DNA repair enzyme uracil DNA glycosylase (UDG) is to use DNA substrates in which the number of hydrogen bonds between the target uracil (U) and its opposing base (X) are systematically varied (Figure 2) (14). Our original hypothesis was that removal of base pair hydrogen bonds would thermodynamically facilitate uracil flipping and enzyme binding by destabilizing the uracil in the DNA base stack (i.e., by raising the free energy of the lower left corner of the reaction coordinate diagram in Figure 1B). A LFEP analysis has confirmed this initial expectation, where a strong linear correlation was found between the free energy of UDG binding to a series of DNA duplexes with increasingly destabilized $U \cdot X$ base pairs ($m = -0.37$, $R^2 = 0.86$) (14). A key question arising from this thermodynamic study is the mechanistic origin of the dramatically increased binding affinity to destabilized damaged sites.

Here we use a similar LFEP approach to correlate the effects of stepwise ablation of hydrogen bonds in the $U \cdot X$

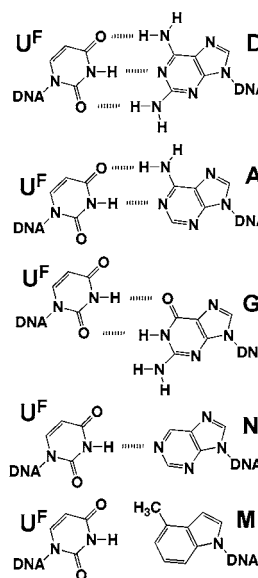
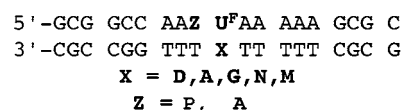


FIGURE 2: Sequences of 19-mer DNA substrates used in the kinetic studies and hydrogen-bonding structures of the $U^F \cdot X$ base pairs based on the literature: $U^F \cdot D$ (57, 58), $U^F \cdot G$ (59), $U^F \cdot N$ (47, 48), and $U^F \cdot M$ (29). Abbreviations: M, 4-methylindole nucleotide; D, dianinopurine nucleotide; N, purine nucleotide; U^F, 2'- β -fluoro-2'-deoxyuridine nucleotide.

base pair with the rate of base extrusion (nucleotide rotation reaction coordinate) and the rate of conformational changes in the DNA (bending reaction coordinate). To monitor the progress along these coordinates, we employ two indirect signal changes. The first follows the increase in fluorescence of a 2-aminopurine (2-AP) nucleotide adjacent to the target uracil (5). This signal is sensitive to the extent of stacking

¹ Abbreviations: LFEP, linear free energy perturbation; 2-AP, 2-aminopurine; FRET, fluorescence resonance energy transfer; M, 4-methylindole nucleotide; D, dianinopurine nucleotide; Y, pyrene nucleotide; N, purine nucleotide; Φ , tetrahydrofuran abasic nucleotide; U^F, 2'- β -fluoro-2'-deoxyuridine nucleotide; UDG, uracil DNA glycosylase.

of 2-AP with the uracil (15, 16) and thus reports on conformational changes in the DNA as well as progress along the base rotation coordinate (6, 17–19). The second signal involves a decrease in UDG tryptophan fluorescence that accompanies a conformational change in UDG as it closes around the fully extrahelical uracil. This signal change has been shown by structural (8), kinetic (5, 6, 17, 20), and mutagenesis studies (6, 17) to correlate exclusively with formation of the final hydrogen-bonding and stacking interactions of the uracil base within the enzyme active site (i.e., formation of the final extrahelical state depicted in the upper right corner of Figure 1B). These studies suggest a base flipping pathway in which DNA bending precedes base rotation, thereby opening an unhindered passage by which uracil may exit the duplex. Surprisingly, binding studies of both rigid and flexible DNA duplexes indicate that a large amount of enzyme binding energy is consumed during the process of DNA bending. This finding appears to be a general feature of base flipping enzymes and is consistent with previous suggestions based on structural and biochemical observations with UDG as well as other base flipping enzymes (2, 21–27). Accordingly, strong binding of UDG to flexible target sites results not only from weakened hydrogen bonding of the uracil but also because flexible sites require less enzyme binding energy to bend. These unique features of target site recognition by UDG may be shared by other base flipping enzymes (28).

EXPERIMENTAL PROCEDURES

Materials. The 2'-deoxynucleoside phosphoramidites, CPG supports, and DNA synthesis reagents were purchased from Glen Research (Sterling, VA), except for 2'- β -fluoro-2'-deoxyuridine (U^F), 4-methylindole nucleoside (M), and pyrene nucleoside (Y), which were synthesized as described previously (5, 29, 30). The oligonucleotides were synthesized using standard phosphoramidite chemistry on an Applied Biosystems 392 synthesizer. The oligonucleotides were purified by anion-exchange HPLC (Zorbax Oligo), followed by C-18 reversed-phase HPLC (Phenomenex Aqua column). Fractions containing pure oligonucleotide were concentrated to dryness in vacuo, redissolved in MilliQ water, and stored at -20°C . The purity of the oligonucleotides was assessed by matrix-assisted laser desorption mass spectroscopy and denaturing polyacrylamide gel electrophoresis. The concentration of each oligonucleotide was determined using its extinction coefficient at 260 nm (31). DNA duplexes were hybridized in 10 mM Tris-HCl (pH 8.0) and 60 mM NaCl as described previously (5). The purification of *Escherichia coli* UDG has been described previously (32).

K_D Measurements. The K_D values for binding of the $U^F\cdot X$ duplexes to UDG were measured using a kinetic competitive inhibition fluorescence assay under conditions where the apparent K_i is equal to the K_D value (i.e., $[S] \ll K_m$, where S exhibits rapid equilibrium binding). Reaction mixtures (148.5 μL) containing 10 mM Tris-HCl (pH 8.0), 60 mM NaCl, 12.5 $\mu\text{g/mL}$ BSA, 2 μM AUPA substrate, and a variable amount of the $U^F\cdot X$ duplex were placed in a 0.3 cm quartz cuvette, and its fluorescence emission at 370 nm was monitored every 10 s at 25°C on a SPEX FluoroMax-3 fluorometer ($\lambda_{\text{ex}} = 320\text{ nm}$) until the signal stabilized. UDG (1.5 μL) was then added to the reaction mixture to give a final concentration of 0.25 nM. The progress of the reaction

was then monitored as described above for 5 min. The K_D for each duplex was determined by fitting to the equation

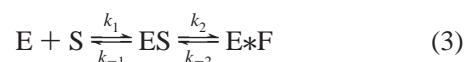
$$k_i/k_0 = 1/(1 + [U^F\cdot X]/K_D) \quad (1)$$

where k_i is the inhibited rate and k_0 is the rate in the absence of competitor DNA.

Stopped-Flow Fluorescence. The observed rate constants for the formation of UDG•DNA complexes were measured on an Applied Photophysics 720 stopped-flow fluorometer (Surrey, U.K.). All measurements were performed under pseudo-first-order conditions where the concentration of the unlabeled component was at least 4-fold greater the concentration of the labeled species. All measurements were made using a buffer containing 10 mM Tris-HCl (pH 8.0), 60 mM NaCl, and 1 mM DTT. For experiments where changes in 2-AP fluorescence were observed, excitation was at 315 nm, and a 360 nm long-pass emission filter was used. In experiments where tryptophan fluorescence was measured, an excitation wavelength of 290 nm was used with a 335 nm long-pass emission filter. All kinetic traces were well fitted to a first-order rate expression to obtain the observed rate constant (k_{obsd} , eq 2).

$$F_t = \Delta F \exp(1 - k_{\text{obsd}}t) + F_0 \quad (2)$$

For the experiments in which the rate of 2-AP signal increase of the DNA was monitored, a solution containing a variable amount of UDG was rapidly mixed with a fixed concentration of 2-AP-labeled DNA. Final concentrations of enzyme and DNA after mixing were in the range 0.4–6.4 μM and 100 nM, respectively. For the experiments in which the rate of tryptophan signal decrease of the enzyme was monitored, a solution containing a variable amount of DNA was rapidly mixed with a fixed concentration of the enzyme. In these experiments, the final concentrations of DNA and enzyme after mixing were in the range 0.4–12.8 μM and 100 nM, respectively. Measurements with the $U^F\cdot M$ duplex using 2-AP fluorescence could not be accurately made above 1.6 μM because only 20% of the observable signal remained after the ~ 1 ms dead time of the instrument had elapsed. As previously observed, plots of k_{obsd} against $[\text{UDG}]$ or $[U^F\cdot X]$ were hyperbolic, indicating a multistep binding mechanism. The kinetic parameters were extracted by fitting the data to a minimal two-step binding model (eqs 3 and 4),



$$k_{\text{obsd}} = \frac{k_{\text{off}}[S] + k_{\text{off}}}{K'[S] + 1} \quad (4)$$

where $K' = k_1/(k_{-1} + k_{\text{max}})$ is the apparent affinity constant, $k_{\text{on}} = K'(k_2 + k_{-2})$ is the apparent second-order association rate constant, and $k_{\text{off}} = k_{-1}k_{-2}/(k_{-1} + k_2 + k_{-2})$ is the overall dissociation rate constant. The maximum rate constant for the unimolecular rearrangements detected by the 2-AP and tryptophan fluorescence signals is measured by the asymptote $k_{\text{max}}^{2\text{AP}}$ (or $k_{\text{max}}^{\text{Trp}} = k_2 + k_{-2}$) (5). Although base flipping has been previously shown to involve two internal steps rather than the one shown in eq 3 (6, 17, 20), the simplified analytical expression of eq 4 is very useful for comparing

the kinetic behavior of a series of substrates or mutated enzymes (6).

Very slow dissociation rate constants determined from the concentration dependence of k_{obsd} (eq 3) have a high degree of uncertainty because k_{off} is derived from the y-intercept of a plot of k_{obsd} against [S]. Under such conditions, a trapping experiment is required to obtain a more precise value for k_{off} . Trapping experiments for the substrates containing D, A, G, and N in the base pair with uracil were performed by rapidly mixing a solution consisting of UDG (200–500 nM) and 2-AP-labeled DNA (200 nM) with a large molar excess of nonfluorescent trapping duplex (20 μ M AU^FA-11) (6) using a stopped-flow fluorometer. The time-dependent decrease in 2-AP fluorescence as the bound DNA irreversibly dissociated was fit to a single-exponential decay to obtain k_{off} (eq 5). Because of the very slow dissociation of U^F•M,

$$F_t = \Delta F \exp(-k_{\text{off}}t) + F_0 \quad (5)$$

its k_{off} was measured by manually mixing a solution consisting of UDG (100 nM) and PU^F•M (100 nM) with 1 μ M high-affinity trap, AU^FA/TMT-15 (14). This reaction was performed using a 0.3 cm path length quartz cuvette (150 μ L) and a Spex FluoroMax-3 fluorometer. The time-dependent decrease in 2-AP emission at 370 nm was followed with excitation at 315 nm.

RESULTS

Binding of UDG to Destabilized Damaged Sites. We synthesized a series of 19-mer DNA duplexes containing a series of U^F•X base pairs using a sequence based on a duplex previously used in rapid kinetic studies of base flipping by UDG (Figure 2) (5). The 2' fluorinated deoxyuridine substrate analogue (U^F) is an extremely slow substrate for UDG ($t_{1/2} \sim 1$ day), allowing measurements of DNA binding and base flipping without the complication of glycosidic bond cleavage (5). The affinity of UDG for these duplexes was measured using a competitive inhibition kinetic assay (14). As previously observed for a series of 15-mer duplexes with a different sequence (14), the K_D values for the 19-mers decreased incrementally as the U^F•X base pair weakened (Figure 3A and Table 1). The U^F•M duplex, which contains no hydrogen bonds, binds a striking ~ 43400 -fold (6.3 kcal/mol) more strongly than the duplex containing the U^F•D base pair with three hydrogen bonds. Of note, the incremental decrease in the free energy of binding was only about 1 kcal/mol when the first and second hydrogen bonds were removed from the U^F•D base pair (Figure 3B), which is similar to estimates of the free energy contribution of individual hydrogen bonds to the stability of duplex nucleic acids (33). In contrast, removal of the last hydrogen bond to form the U^F•M pair resulted in a much larger 4 kcal/mol decrease in binding free energy, indicating the presence of additional energetic contributions. These additional contributions may involve a loss of both hydrogen bond and base stacking interactions, leading to a large increase in conformational entropy of the U^F•M base pair (14). This result suggests that, with respect to the free energy of binding, increased flexibility of the site may play a more important role than the enthalpic benefit of removing single hydrogen bonds (see below).

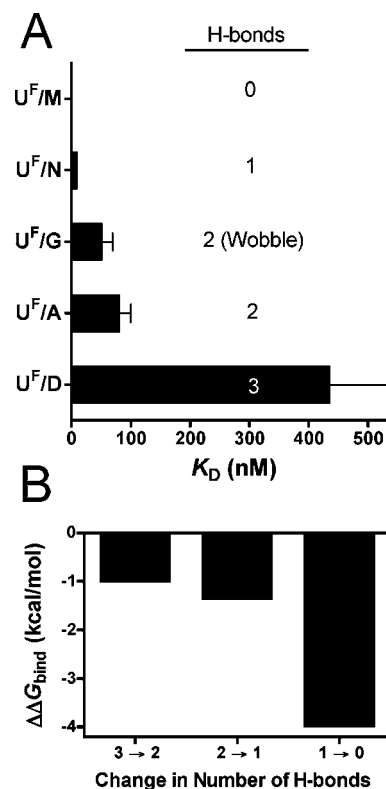


Figure 3

FIGURE 3: Binding affinities of UDG for U^F•X DNA duplexes and incremental change in binding free energy when each hydrogen bond is removed. (A) Trend in dissociation binding constants as base pair hydrogen bonds are removed (see also Table 1). The number of hydrogen bonds in the U^F•X base pair is shown to the right of the bars. (B) Incremental change in binding free energy as each hydrogen bond is removed from the U^F•X base pair ($\Delta\Delta G_{\text{bind}} = -RT \ln K_D^n/K_D^{n-1}$, where n is the number of hydrogen bonds in the base pair).

Binding Kinetics. To dissect the origins of the dramatically enhanced binding affinity of the duplexes with destabilized U^F•X base pairs, the kinetics of association and dissociation were measured. Association rates were measured using stopped-flow fluorescence measurements by monitoring either the increase in 2-AP fluorescence that accompanies uracil unstacking and DNA bending (Figure 4A) or the decrease in tryptophan fluorescence that marks the attainment of the final extrahelical state (Figure 4B) (5, 6, 17). As observed in previous studies using duplexes with U^F•A and U^F•G base pairs (5), plots of k_{obsd} versus concentration were hyperbolic when either the 2-AP or tryptophan fluorescence signals were monitored (panels A and B of Figure 5, respectively). This kinetic behavior indicates a change in rate-limiting step from bimolecular encounter at low concentrations of the varied reactant to a unimolecular conformational change of the DNA and enzyme at high reactant concentrations. All hyperbolic plots were fit to the two-step binding model (eq 4), and the kinetic constants obtained from this analysis are reported in Table 1.

There are revealing aspects of the kinetics for the 2-AP and tryptophan fluorescence changes. First, for the substrate with the most stable base pair (U^F•D), the association kinetics measured using the 2-AP signal ($k_{\text{on}}^{2\text{AP}}$) is 3-fold faster than when the tryptophan signal is followed ($k_{\text{on}}^{\text{Trp}}$) (Table 1).

Table 1: Binding Affinities and Kinetic Parameters for UDG Association and Dissociation with U^F·X Duplexes

duplex	K_D (nM)	K'^{2AP} (μM^{-1})	K'^{Trp} (μM^{-1})	k_{on}^{2AP} ($\mu\text{M}^{-1}\cdot\text{s}^{-1}$)	$k_{\text{on}}^{\text{Trp}}$ ($\mu\text{M}^{-1}\cdot\text{s}^{-1}$)	k_{off} (s^{-1})	K_D^{calc} (nM) ^d
U ^F ·D	434 ± 100 ^a	0.29 ± 0.05	0.12 ± 0.02	206 ± 22	72 ± 7	44 ± 1.3	214 ± 24
U ^F ·A	80 ± 20 ^{a,b}	0.08 ± 0.01	0.13 ± 0.01	236 ± 10	172 ± 8	26 ± 0.3	110 ± 4
U ^F ·G	50 ± 20 ^{a,b}	0.10 ± 0.01	0.25 ± 0.08	292 ± 13	331 ± 56	6.9 ± 0.04	24 ± 1
U ^F ·N	8.2 ± 0.3 ^c	0.20 ± 0.04	0.14 ± 0.01	282 ± 28	315 ± 5	6.0 ± 0.05	21 ± 2 ^c
U ^F ·M	0.01 ± 0.002 ^c	0.32 ± 0.15	0.64 ± 0.08	292 ± 40	666 ± 45	0.018 ± 0.001	0.062 ± 0.009 ^c

^a K_D was determined by directly monitoring binding of DNA to UDG using 2-AP fluorescence (see Experimental Procedures). ^b Values previously reported (5). ^c K_D was determined using competitive inhibition assay with AU^F·X substrate (see Experimental Procedures). The affinity of AU^F·A is 2-fold greater than that of PU^F·A (17). ^d K_D calculated from $k_{\text{off}}/k_{\text{on}}^{2AP}$.

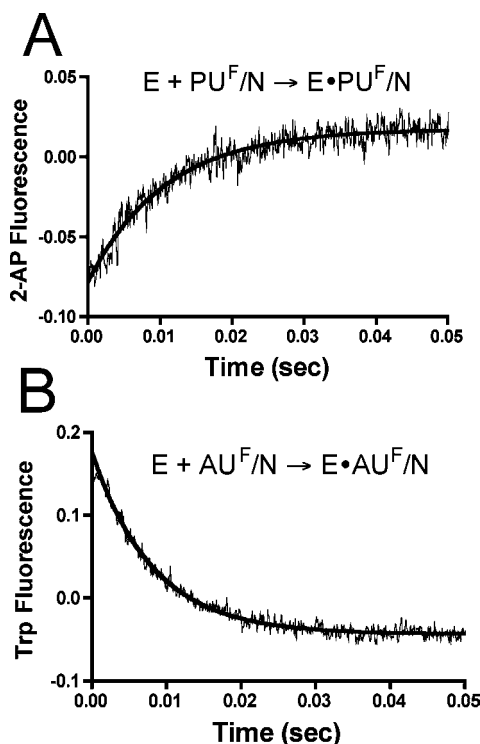


FIGURE 4: Stopped-flow fluorescence kinetic measurements of UDG association with U^F·X duplex DNA. (A) Approach to equilibrium association rate of PU^F·N (0.1 μM) and UDG (0.4 μM) monitored by the increase in 2-AP fluorescence. (B) Approach to equilibrium association rate of UDG (0.1 μM) and AU^F·N (0.4 μM) monitored by the decrease in tryptophan fluorescence of UDG.

However, as hydrogen bonds are removed from the base pair k_{on}^{2AP} and $k_{\text{on}}^{\text{Trp}}$ become indistinguishable. This result arises because k_{on}^{2AP} is nearly invariant across the whole series of duplexes, while $k_{\text{on}}^{\text{Trp}}$ increases by about 9-fold (Figure 5C).² The observation that $k_{\text{on}}^{\text{Trp}}$ is slower than k_{on}^{2AP} when the base pair is strong requires that the conformational change in the enzyme lags behind the structural perturbation in the DNA that results in the 2-AP fluorescence increase. The strong dependence of $k_{\text{on}}^{\text{Trp}}$ on base pair hydrogen bonding indicates that these hydrogen bonds are broken very late

² The asymptotic k_{obsd} values for the 2-AP signal ($k_{\text{max}}^{2AP} = k_2^{2AP} + k_{-2}^{2AP}$; see eq 3) increase by about 2-fold between U^F·D and the U^F·A or U^F·G substrates but then decrease from these peak values by about 2-fold for the U^F·N and U^F·M substrates. This behavior may be qualitatively understood using the simplified two-step kinetic analysis we have employed. That is, k_2^{2AP} and k_{-2}^{2AP} are complex rate constants that reflect the interconversion of **I**₁, **I**₂, and **P** on the enzyme: k_{max}^{2AP} first increases as the conversion of **I**₂ → **P** becomes more rapid due to hydrogen bond ablation and then decreases because the reverse rate (k_{-2}^{2AP}) becomes negligible for the U^F·N and U^F·M substrates due to stabilization of **I**₂ and **P** (i.e., $k_{\text{max}}^{2AP} \sim k_2^{2AP}$; see Discussion).

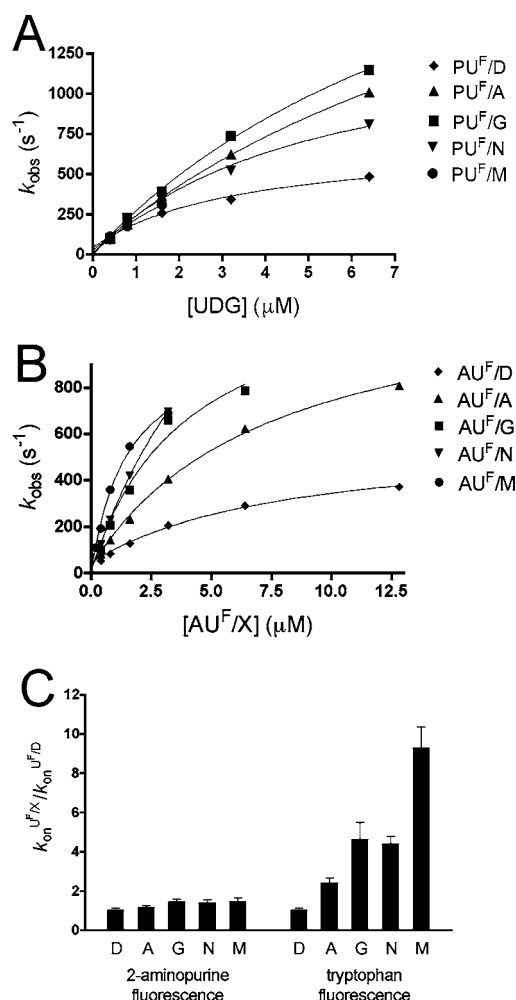


FIGURE 5: Concentration dependence of the apparent association rate constants (k_{obs}) for the five different U^F·X base pairs and relative k_{on} values as compared to the U^F·D substrate. (A) Dependence of k_{obs} on the concentration of UDG followed by changes in 2-AP fluorescence. Data were fit to the two-step kinetic model (eqs 3 and 4). (B) Dependence of k_{obs} on the concentration of DNA followed by changes in Trp fluorescence. Data were fit to the two-step kinetic model (eqs 3 and 4). (C) Effect of changing U^F·X base pair strength on k_{on} , expressed as the ratio of $k_{\text{on}}(\text{U}^{\text{F}}\cdot\text{X})$ to $k_{\text{on}}(\text{U}^{\text{F}}\cdot\text{D})$. Ratios determined from the kinetics of the 2-AP and tryptophan fluorescence changes are displayed separately.

along the reaction coordinate, *after* the step that is reported on by the 2-AP signal and *during* the final conformational step when the enzyme closes around the uracil base. This result implies that the 6-fold 2-AP signal increase reflects the formation of an intermediate (**I**₂, Figure 1) in which the uracil stacking interactions with the adjacent 2-AP base are perturbed, but the uracil retains hydrogen bonding with its

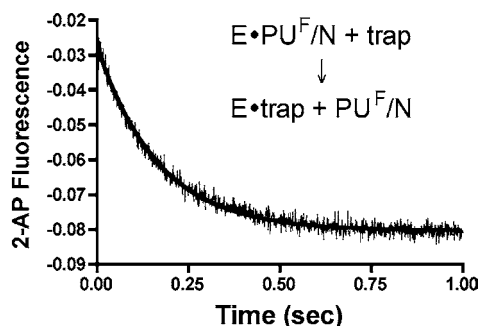


FIGURE 6: Determination of k_{off} using a trapping experiment. A complex of $\text{PU}^{\text{F}}\cdot\text{N}$ ($0.20\ \mu\text{M}$) and UDG ($0.25\ \mu\text{M}$) was rapidly mixed with a large excess of nonfluorescent $\text{AU}^{\text{F}}\text{A-11}$ single-stranded DNA ($20\ \mu\text{M}$) (6). The k_{off} is determined from fitting the time-dependent decrease in the 2-AP fluorescence of $\text{PU}^{\text{F}}\cdot\text{N}$ as it is irreversibly released from UDG (eq 5).

partner base (6, 17). Since the preceding intermediate (I_1 , Figure 1) is not detectably bent (7), and the subsequent extrahelical state is bent by about 40° (8), then the structural implication of the increase in 2-AP fluorescence is that the DNA is bent in the transition state leading to I_2 ,³ although the magnitude of the bending at this point along the reaction coordinate is not known. It is important to point out that the 9-fold increase in k_{on} pales in comparison with the approximately 43000-fold increase in UDG binding affinity when these three hydrogen bonds are ablated. Thus, the vast majority of the binding affinity increase is not attributable to k_{on} .

The above results strongly suggest that the large thermodynamic benefit of removing hydrogen bonds in the $\text{U}^{\text{F}}\cdot\text{X}$ base pair arises from profound differences in the dissociation rates of these duplexes. To establish this point, the dissociation rates (k_{off}) were measured using an irreversible trapping experiment (5) in which a solution of $\text{UDG}\cdot(\text{PU}^{\text{F}}\cdot\text{X})$ was rapidly mixed with a large excess of unlabeled trapping DNA, and the decrease in the 2-AP signal was followed (Figure 6). The dissociation rate constants are reported in Table 1 and are consistent with the y-intercepts (k_{off}) estimated from the concentration dependence of the association rates in Figure 5A,B. The dissociation rates were strongly dependent on the nature of the $\text{U}^{\text{F}}\cdot\text{X}$ base pair, decreasing by over 4 orders of magnitude over the entire series, with $k_{\text{off}}(\text{U}^{\text{F}}\cdot\text{D}) = 44\ \text{s}^{-1}$ and $k_{\text{off}}(\text{U}^{\text{F}}\cdot\text{M}) = 0.018\ \text{s}^{-1}$. These data provide strong evidence that removal of target base pair hydrogen bonds leads to profound stabilization of one or more enzyme-bound species on the uracil flipping pathway. Since the kinetic effect of removing hydrogen bonds appears very late during the uracil flipping process (i.e., during the

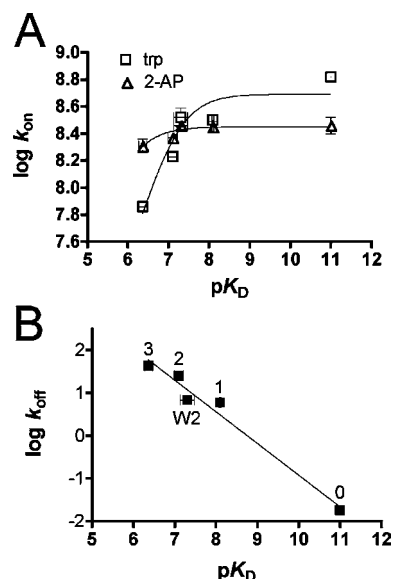


FIGURE 7: Logarithmic dependences of k_{on} and k_{off} on the binding affinity of UDG for $\text{U}^{\text{F}}\cdot\text{X}$ DNA (pK_{D}). (A) Logarithmic dependence of k_{on} on pK_{D} . The data were fit to the empirical equation $\log k_{\text{on}} = \log k_{\text{diff}} - \log(1 + 10^{\text{pK}_0 - \text{pK}_{\text{D}}})$, where k_{diff} is the rate constant for diffusional encounter and pK_0 is an empirical constant. (B) Linear dependence of $\log k_{\text{off}}$ on binding affinity (slope = 0.74 ± 0.07 , $R^2 = 0.97$). For reference, the number of hydrogen bonds in each destabilized $\text{U}^{\text{F}}\cdot\text{X}$ base pair is indicated above each data point.

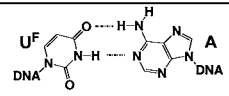
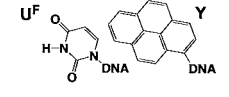
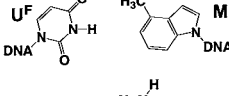
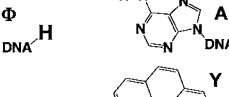
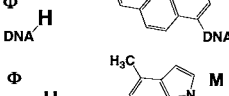
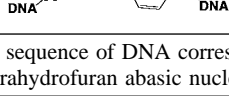
last conformational step monitored by tryptophan fluorescence; see above), then the stabilized species must be one or both of the complexes shown in the upper and lower right-hand corners of Figure 1B (I_2 , P).

Linear Free Energy Correlations. The quantitative dependence of k_{on} and k_{off} on UDG binding affinity provides further insights into the mechanistic basis for the enhanced binding of duplexes with destabilized base pairs. A plot of $\log k_{\text{on}}^{2\text{AP}}$ versus pK_{D} is nearly flat across the whole series (Figure 7A, triangles) with an average value $k_{\text{on}} = 2.8 \times 10^8\ \text{M}^{-1}\ \text{s}^{-1}$ indicative of near diffusion-controlled binding except for the substrate with the most stable base pair ($\text{U}^{\text{F}}\cdot\text{D}$) which falls slightly below this average value. Thus, all internal kinetic steps preceding and including the 2-AP fluorescence change are rapid as compared to the encounter rate. In contrast, $k_{\text{on}}^{\text{Trp}}$ shows a linear increase as a function of pK_{D} for the more stable $\text{U}^{\text{F}}\cdot\text{D}$ and $\text{U}^{\text{F}}\cdot\text{A}$ duplexes, followed by downward curvature to a plateau value of $k_{\text{on}}^{\text{Trp}} = 4 \times 10^8\ \text{M}^{-1}\ \text{s}^{-1}$ for the most destabilized base pairs (Figure 7A, squares). This behavior indicates a change in rate-limiting step as a function of base pair stability from the step giving rise to the tryptophan fluorescence change to that involving diffusion-controlled encounter. We conclude that the thermodynamic stability of the $\text{U}^{\text{F}}\cdot\text{X}$ base selectively alters the activation barrier for the step reported by $k_{\text{on}}^{\text{Trp}}$.

In contrast with $\log k_{\text{on}}$, there is a linear dependence of $\log k_{\text{off}}$ on pK_{D} , with the dissociation rate decreasing steeply as UDG affinity increases (Figure 7B). From the slope of this correlation (-0.74 ± 0.07), it can be inferred that about 74% of the difference in binding energies between any two $\text{U}^{\text{F}}\cdot\text{X}$ duplexes arises from k_{off} . The simplest mechanism that accounts for the small effect of base pair strength on k_{on} , and the large effect on k_{off} , is the progressive stabilization of one or more enzyme-bound species as hydrogen bonds are removed (see Discussion).

³ We were unable to detect DNA bending by UDG employing standard fluorescence resonance energy transfer (FRET) methods and DNA substrates with fluorescence donor and acceptor groups on each end (51, 52). We attribute this result to interaction of the rhodamine fluorophore with binding sites in the free DNA (53, 54), which may be disrupted when the enzyme binds, resulting in no FRET change. Based on crystallographic models of bent DNA bound to UDG (not shown), we would have expected a substantial 1.4-fold FRET increase. The same negative result was obtained when the donor or acceptor was placed on the 3' or 5' end of the uracil-containing strand or when a 15-mer or 19-mer duplex was used. We note that an unexpected and small 15% FRET decrease was recently reported for *EcoRI* methyltransferase which also flips a DNA base by a putative bending mechanism, although an increase was expected (55). We conclude that FRET results should be interpreted cautiously in base flipping systems.

Table 2: Effect of Pyrene Nucleoside on Binding Affinity of UDG to Uracil and the Abasic Site Containing DNA^a

Duplex DNA	K_D (nM)
	300 ± 50
	120 ± 30
	0.24 ± 0.03
	100 ± 10
	8700 ± 675
	18 ± 3.0

^a The sequence of DNA corresponds to 15-mer duplexes in ref 14. Φ = tetrahydrofuran abasic nucleotide analogue.

Base Pair Hydrogen Bonds, DNA Flexibility, and UDG Binding. To test the role of DNA flexibility on binding affinity, we employed a novel structure–reactivity approach. This strategy takes advantage of the decrease in flexibility of the target site when it is modified with a rigid pyrene nucleotide on one strand (Table 2) (17, 30, 34–36). We originally used a uracil–pyrene base pair ($U^F \cdot Y$) to preorganize uracil in an extrahelical conformation (34), with the rationale that the planar aromatic ring structure of pyrene would fully encompass the entire volume normally occupied by the standard $U^F \cdot A$ base pair, thereby presenting the uracil to the enzyme in an extrahelical conformation that favors binding (i.e., the energetic penalty for uracil flipping is prepaid, allowing for tighter binding). In accord with this reasoning, DNA that contained a $U^F \cdot Y$ pair fully rescued the catalytic and base flipping activities of UDG mutants that were incapable of pushing uracil from the DNA base stack in the context of a $U^F \cdot A$ base pair (17, 34).

Although base flipping mutants showed dramatic increases in their site-specific binding affinity and catalytic activity when a $U^F \cdot Y$ pair was used, wild-type UDG bound with nearly equal affinity to DNA containing either a $U^F \cdot Y$ or $U^F \cdot A$ base pair (34). This finding is reaffirmed in Table 2 where the binding affinity of UDG to a 15-mer duplex containing a $U^F \cdot A$ and $U^F \cdot Y$ pair is shown to differ by only 2.5-fold. This result is in dramatic contrast to the 1200-fold tighter binding of the $U^F \cdot M$ duplex as compared to $U^F \cdot A$ using the same 15-mer context (Table 2). These observations indicate that there is a cryptic energetic penalty for binding of the $U^F \cdot Y$ duplex that is not present for binding of the $U^F \cdot M$ duplex even though $U^F \cdot M$ and $U^F \cdot Y$ both lack hydrogen bonds to the uracil base. Thus a key question is the basis for the dramatically tighter binding of the $U^F \cdot M$ duplex as compared to $U^F \cdot Y$ ($\Delta\Delta G = 3.7$ kcal/mol).

A likely explanation is that the pyrene nucleoside serves to stiffen the DNA at the apex of the bend induced during base flipping and thus requires a greater expenditure of

binding energy by UDG in order to distort the DNA as compared to $U^F \cdot M$. This idea is consistent with previous computational findings suggesting that decreased DNA flexibility can lead to reduced UDG binding affinity by increasing the resistance of the DNA to adopt the bend enforced by protein binding (10). Increased structural rigidity introduced by pyrene is chemically reasonable because the extended π aromatic system of the pyrene base pair mimic is much more rigid than the native $U^F \cdot A$ base pair which is only restrained by noncovalent hydrogen bonding. Assuming this interpretation for binding of the $U^F \cdot Y$ base pair, its unfavorable rigidity is offset by the energetic benefit of extrahelical preorganization of the uracil. Thus, the binding affinity of $U^F \cdot Y$ to UDG is similar to a $U^F \cdot A$ base pair because of compensatory energetic effects, but $U^F \cdot M$ binds more tightly because of the increased flexibility of the site (Table 2).

To further test the proposal that rigidity of the DNA has an adverse effect on UDG binding affinity, we tested the effect of the opposing base (A, M, or Y) on binding of DNA that contained an abasic site product analogue (Φ) (Table 2). Structural studies have shown that UDG can bind specifically to and flip out abasic sites to form a structurally indistinguishable complex as compared to substrate analogue DNA (2, 8, 37, 38). However, because there is no requirement for rupture of a base pair upon flipping of an abasic nucleotide, and there is no uracil attached to form hydrogen bonds and stacking interactions in the UDG active site, the changes in binding affinity between the $\Phi \cdot Y$ and $\Phi \cdot M$ duplexes should largely reflect differences in the intrinsic flexibility of the target site. As shown in Table 2, $\Phi \cdot M$ binds ~ 480 -fold more tightly than $\Phi \cdot Y$ and 5-fold more tightly than $\Phi \cdot A$. The observed greater affinity of UDG for $\Phi \cdot M$ over $\Phi \cdot Y$ is identical to that of $U^F \cdot M$ and $U^F \cdot Y$ (Table 2). We conclude from these results that pyrene introduces an unfavorable free energy contribution to binding of as much as 3.7 kcal/mol arising from its increased rigidity and stacking as compared to sites that contain A or M as the opposing base. In conclusion, the observation that flexible abasic sites bind with much higher affinity than rigid abasic sites provides strong experimental evidence that specific recognition involves a flexibility component regardless of the presence of a uracil base.

DISCUSSION

The mechanism by which enzymes obtain extraordinary specificity for their respective target sites in a large background of random DNA sequences has long been an active area of research. In this general area, DNA glycosylases have evolved a unique solution to this problem that is divergent from the sequence-dependent mechanisms of restriction enzymes, transcription factors, and repressor proteins. Obtaining a fundamental understanding of the kinetic and thermodynamic origins of DNA repair glycosylase specificity has the potential to allow rational targeting of these enzymes to engineered DNA sites that display features that promote enhanced binding and reactivity. Indeed, we have previously shown that simple rules derived from mechanistic studies of base flipping and catalysis can be used to alter UDG's specificity to recognize a cytosine base opposite to pyrene (i.e., a $C \cdot Y$ base pair) rather than uracil (39). Such findings offer the promise of altering the coding sequences of genes

in vivo using engineered glycosylases in combination with assisting targeting molecules such as pyrene-containing oligonucleotides.

The Reaction Coordinate for Base Flipping. It is difficult to imagine the pathway for flipping a base 180° from the DNA base stack ultimately leading to its precise docking into an enzyme active site. It is equally difficult to view how an enzyme could form specific interactions to guide a target base through a trajectory that arcs widely out of the DNA helix (Figure 1A) (7, 35). The present data suggest plausible mechanistic answers to these questions that are consistent with previous NMR dynamic and structural studies of the earliest steps in the base flipping process (7, 35) and also crystallographic studies of the final extrahelical complex (8).

To understand how an enzyme might facilitate the overall base flipping process, it is useful to consider the intrinsic kinetic and thermodynamic problems that must be overcome in extruding a base from the B DNA base stack (Figure 8). First, NMR imino proton exchange experiments have established that T•A, T•G, and U•A base pairs rapidly open at room temperature (7, 40, 41). These opening rates are greater than or equal to the rate constants for kinetic steps on the base flipping of UDG (Table 1), indicating that spontaneous base pair opening provides kinetically competent motions for seeding the enzymatic base flipping process. Indeed, recent studies of imino proton exchange of UDG-bound DNA have established that UDG does not alter the opening rate but instead substantially slows the closing rate constant of open T•A base pairs (7). This result is intuitively satisfying because the base pair closing rates in B DNA are exceedingly fast ($\sim 10^7$ s⁻¹) (42, 43), and one major problem in facilitating the migration of a base forward along the flipping trajectory is to prevent its retrograde motion back into the DNA base stack (Figure 8A). On the basis that adjacent T•A and G•C base pairs showed identical closing rates in the UDG complex ($\sim 10^5$ s⁻¹), but these rates differed by 25-fold in the free DNA, we proposed that UDG uses nonspecific DNA backbone interactions to increase the lifetime of a high-energy open state of the bound DNA that occurs very early on the flipping trajectory (**I**₁, Figure 8A) (7). This open state of UDG-bound DNA is still extremely unstable and exists in dynamic equilibrium with closed B DNA. Nevertheless, the equilibrium is now shifted more toward the open state as compared to the free DNA (Figure 8A).

The next intermediate (**I**₂) on the flipping pathway has an altered B DNA structure such that the uracil stacking interactions with the adjacent 2-AP are perturbed, but on average, the base pair hydrogen bonds are still intact (Figure 8A). The experimental observations that support these features of **I**₂ are (i) the rapid 2-AP fluorescence change indicating a perturbation in the stacking of uracil with 2-AP during the formation of **I**₂ (Figure 4A) and (ii) that the kinetic step between **I**₂ and the final extrahelical state (**P**) is kinetically enhanced as hydrogen bonds are removed from the U•X base pair (Figure 5C), indicating that these bonds persist after formation of **I**₂. Once the final state is attained and the uracil is docked and held securely in the active site by stacking and hydrogen-bonding interactions (8, 38, 44–46), the extrahelical base cannot reenter the DNA stack without climbing the reverse activation barrier back to **I**₂. Thus, the base flipping process may be viewed as the progressive formation of enzyme-bound states in which the

uracil base spends an ever increasing time in an extrahelical conformation as depicted in Figure 8A.

The Importance of Being Flexible. The kinetic and thermodynamic data indicate that as the base pair hydrogen bonds are removed, the second intermediate (**I**₂), or the final extrahelical state (**P**), is increasingly stabilized (Figure 8B). The results also reveal that removal of the first two hydrogen bonds alters binding affinity by ~ 1 kcal/mol per hydrogen bond but that removal of the final hydrogen bond results in a significantly larger 4 kcal/mol increase in the binding affinity (Figure 3B). Assuming that the base pair remains largely intact when the first two hydrogen bonds are broken, which is supported by both NMR and potassium permanganate oxidation sensitivity measurements (14, 47, 48), then the 1 kcal/mol free energy increment may provide an estimate of the individual hydrogen bond enthalpy (33). The free energy change when the last bond is removed is substantially greater than estimates of DNA base hydrogen bond energy (33, 49), indicating the loss of multiple energetic interactions, perhaps including stacking of the uracil with adjacent bases. We suggest that one ramification of total ablation of hydrogen bonds in the U^F•X pair is increased conformational flexibility of the target site. Such conformational flexibility would be expected to facilitate DNA bending and, therefore, binding affinity. The conclusion that hydrogen bond ablation increases target site flexibility is further supported by the 3.7 kcal/mol unfavorable effect of the rigid pyrene (Y) nucleotide on binding as compared to 4-methylindole (M) when these nucleotides are placed opposite to uracil or an abasic site in duplex DNA (Table 2). We infer that pyrene, by forming strong stacking interactions and occupying a space that spans the width of the DNA duplex, significantly increases the resistance of the DNA toward bending, which is in turn reflected in the binding affinity measurements. Of course, the other destabilized base pairs may also exhibit changes in stacking energetics, but these differences are expected to be small given the very conservative changes in the nature of the opposing base (50). These differences in stacking energies, if present, would also be manifested in the flexibility of the DNA and would therefore contribute to the energetics of enzyme-induced DNA bending.

A free energy diagram depicting the significant effect of target site flexibility on binding is shown in Figure 8B. The diagram highlights how weaker U^F•X base pairs are easier to bend due to their increased conformational flexibility, leading to a decrease in the free energy of the bent DNA, reflected in the stability of **I**₂ and **P** in Figure 8B.⁴ In addition to the energetic benefit of increased flexibility of the U^F•M substrate (~ 3 kcal/mol), the **P** state realizes the additional energetic benefit of forming hydrogen bonds with the uracil base without the penalty of breaking hydrogen bonds in the base pair (~ 3 kcal/mol).

The data uncover that a significant portion of UDG's intrinsic binding energy is used to drive the unfavorable process of DNA bending. Thus, a key question is the benefit

⁴ This is similar to our previous conclusion that enthalpically destabilizing the uracil base pair promotes enhanced binding by increasing the probability of extrahelical states that are recognized by the enzyme (14). However, our previous interpretation did not include the effect of base pair flexibility on the energetics of DNA bending. Indeed, increased flexibility appears to be more important than the enthalpic benefit of removing hydrogen bonds.

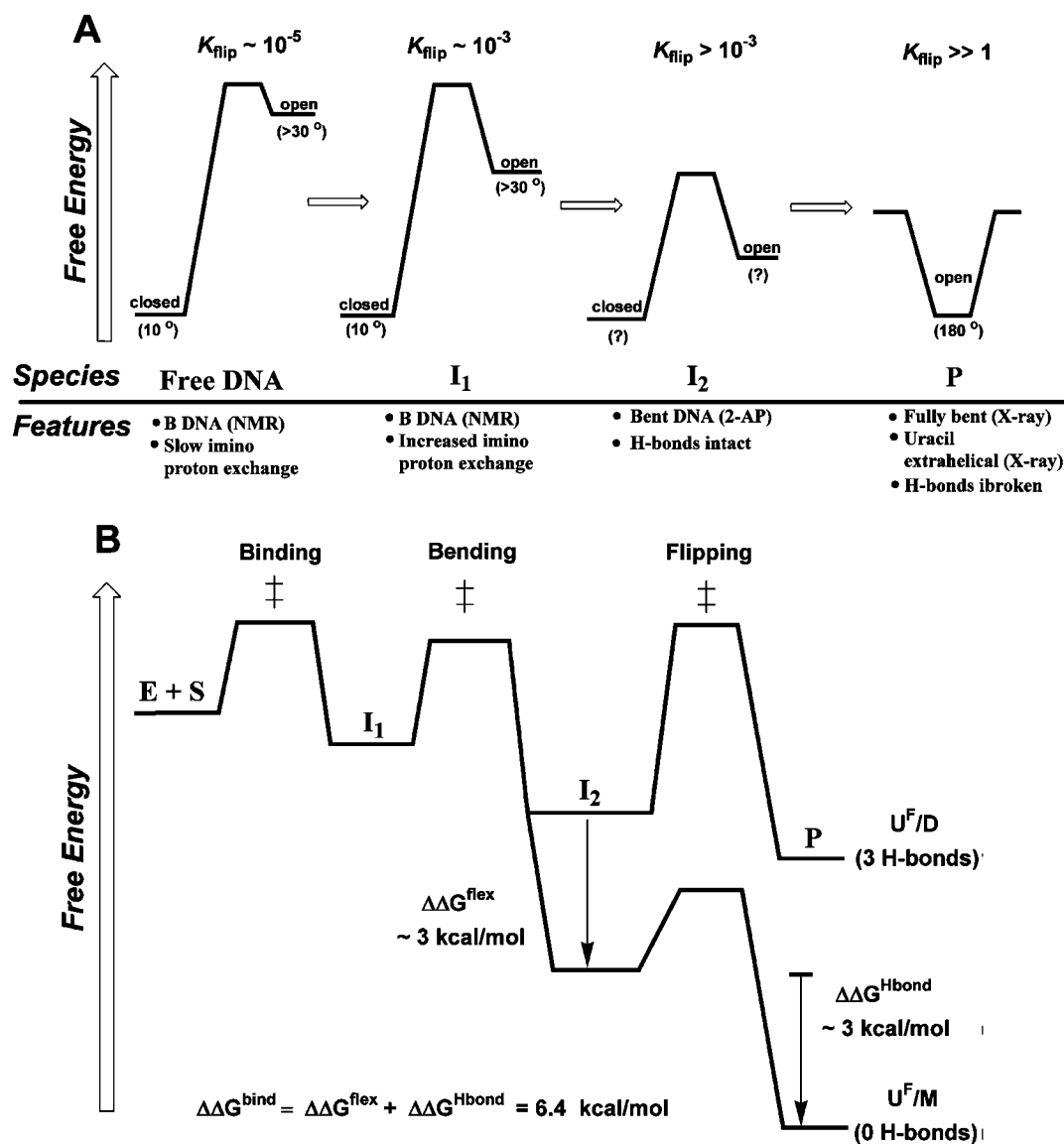


FIGURE 8: Free energy diagrams depicting the increasing base flipping equilibrium in each intermediate along the reaction coordinate for uracil flipping and the effect of base pair destabilization on binding affinity of UDG. (A) Free energy diagrams depicting the increasing equilibrium constant for uracil flipping along the reaction coordinate. The equilibrium, $K_{\text{flip}} = [\text{open}]/[\text{closed}]$, describes uracil in a closed, hydrogen-bonded state in the DNA duplex and an open state that is further along the reaction coordinate for base flipping. This flipping equilibrium is qualitatively depicted for the free DNA, the two enzyme-bound intermediates in Figure 1B (I_1 , I_2), and the fully extrahelical product of the flipping reaction (P). The known structural and dynamic features of each depicted species are listed below each profile. The equilibrium constant K_{flip} has been measured for the free DNA ($\sim 10^{-5}$) and for DNA bound in I_1 ($\sim 10^{-3}$) using NMR solvent magnetization transfer methods at 10 °C (7), establishing that UDG increases the equilibrium by about 100-fold in this early intermediate. Less is known about the structure of I_2 , but the present results indicate DNA bending and largely intact base pair hydrogen bonds. In the P state, the uracil is held tightly in the active site, and the equilibrium is pushed all the way to the fully open state (180° rotation). Access to the closed state requires passage back over the transition state connecting I_2 and P . The estimated rotation angles for the closed and open states are based on the pseudodihedral reaction coordinate defined by Banavali and MacKerell (11, 14), which ranges from 10° for uracil in B DNA to 180° for full rotation of the base out of the major groove. (B) Free energy effect of base pair destabilization on binding affinity of UDG ($\Delta\Delta G^{\text{bind}} = \Delta\Delta G^{\text{flex}} + \Delta\Delta G^{\text{Hbond}}$). The relative free energy profiles for the stable U^F/D duplex and the flexible U^F/M duplex are depicted. The increased flexibility of the U^F/M substrate decreases the free energy of I_2 because less binding energy is required to bend the DNA at this step. The magnitude of the flexibility benefit is estimated as $\Delta\Delta G^{\text{flex}} \sim 3 \text{ kcal/mol}$ (see text). In addition to the flexibility benefit, the absence of base pair hydrogen bonds in U^F/M as compared to the three hydrogen bonds in U^F/D enhances binding of the P state by $\Delta\Delta G^{\text{Hbond}} \sim 3 \text{ kcal/mol}$ (see Figure 3B and text). This benefit arises because for U^F/M there is no energetic price for breaking these bonds in the base pair. For simplicity we have shown that $\Delta\Delta G^{\text{flex}}$ arises entirely at the I_2 state, but this effect is more likely realized at both the I_2 and P states.

derived from this significant energetic expenditure. Since the rate of spontaneous base pair opening is already fast compared to events on the base flipping pathway, it is mechanistically reasonable to propose that the role of enzyme-induced DNA bending is to promote forward commitment along the pathway rather than to lower activation barriers to flipping (i.e., stabilization of extrahelical intermediates).⁵

UDG likely promotes forward migration of the uracil down the pathway by forming nonspecific interactions with the

⁵ Computational studies on UDG and other enzymes have suggested that DNA bending facilitates base flipping (9–12, 51–55), and a recent computational study on HhaI (cytosine-C5)-methyltransferase indicates that this enzyme acts by forming a web of interactions with several high-energy intermediates along the flipping pathway (56).

DNA backbone which become progressively stronger during the formation of two discrete intermediates (**I**₁ and **I**₂, Figure 1B). In the earliest nonspecific intermediate (**I**₁), NMR imino exchange experiments indicate that UDG increases the opening equilibrium ([open]/[closed]) by almost 100-fold (2.8 kcal/mol), largely by decreasing the closing rate (7). In the subsequent intermediate (**I**₂), the magnitude of the equilibrium is currently unknown, but a further increase is suggested because the uracil has substantially decreased stacking interactions with the adjacent 2-AP base arising from DNA bending. Since the presence of uracil is required to achieve **I**₂, then this intermediate must be considered a specific complex in which the enzyme recognizes some feature of the uracil base. The extent or nature of the discrimination between uracil and other normal DNA bases at this intermediate step is currently not known. In conclusion, our view based on these findings is that UDG facilitates base flipping by utilizing binding energy to alter the DNA structure, thereby progressively shifting the flipping equilibrium toward the open state that is ultimately trapped by specific interactions located within the uracil binding pocket. The specific interactions are formed only very late in the process, and thus, the earlier steps in the flipping pathway are largely driven by nonspecific interactions. A further experimental goal with UDG is to unambiguously identify the amino acid side chains that promote the stepwise transition from the closed to open state.

Conclusion. We have uncovered the origins of high-affinity binding of UDG to an extrahelical uracil base. By discretely altering base pair hydrogen bonding and site flexibility, we have found that high binding affinity is determined more by the flexibility of the site than by the enthalpic benefit of removing base pair hydrogen bonds. Thus a significant contribution to enzymatic base flipping is the unfavorable energetic cost of DNA bending. This cost is estimated in two ways: (i) the stiffening effect of pyrene (3.7 kcal/mol) and (ii) the additional favorable energetic effect on binding of removing the last hydrogen bond from the uracil base pair (~3 kcal/mol). In the future, these simple features of high-affinity target site binding by UDG may be utilized to direct UDG to specific sites in genomic DNA.

REFERENCES

- Stivers, J. T., and Jiang, Y. L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases, *Chem. Rev.* **103**, 2729–2759.
- Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., and Tainer, J. A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA, *EMBO J.* **17**, 5214–5226.
- Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature* **384**, 87–92.
- Stivers, J. T. (2004) Site-specific DNA damage recognition by enzyme-induced base flipping, *Prog. Nucleic Acid Res. Mol. Biol.* **77**, 37–65.
- Stivers, J. T., Pankiewicz, K. W., and Watanabe, K. A. (1999) Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase, *Biochemistry* **38**, 952–963.
- Jiang, Y. L., and Stivers, J. T. (2002) Mutational analysis of the base flipping mechanism of uracil DNA glycosylase, *Biochemistry* **41**, 11236–11247.
- Cao, C., Jiang, Y. L., Stivers, J. T., and Song, F. (2004) Dynamic opening of DNA during the enzymatic search for a damaged base, *Nat. Struct. Mol. Biol.* **11**, 1230–1236.
- Parikh, S. S., Walcher, G., Jones, G. D., Slupphaug, G., Krokan, H. E., Blackburn, G. M., and Tainer, J. A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: Conformational strain promotes catalytic efficiency by coupled stereoelectronic effects, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5083–5088.
- Fuxreiter, M., Luo, N., Jedlovsky, P., Simon, I., and Osman, R. (2002) Role of base flipping in specific recognition of damaged DNA by repair enzymes, *J. Mol. Biol.* **323**, 823–834.
- Seibert, E., Ross, J. B., and Osman, R. (2002) Role of DNA flexibility in sequence-dependent activity of uracil DNA glycosylase, *Biochemistry* **41**, 10976–10984.
- Banavali, N. K., and MacKerell, A. D., Jr. (2002) Free energy and structural pathways of base flipping in a DNA gcgc containing sequence, *J. Mol. Biol.* **319**, 141–160.
- Huang, N., Banavali, N. K., and MacKerell, A. D., Jr. (2003) Protein-facilitated base flipping in DNA by cytosine-5- methyltransferase, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 68–73.
- Jencks, W. P. (1985) A primer for the bema hapothle—an empirical approach to the characterization of changing transition-state structures, *Chem. Rev.* **85**, 511–527.
- Krosky, D. J., Schwarz, F. P., and Stivers, J. T. (2004) Linear free energy correlations for enzymatic base flipping: How do damaged base pairs facilitate specific recognition?, *Biochemistry* **43**, 4188–4195.
- Rachofsky, E. L., Osman, R., and Ross, J. B. (2001) Probing structure and dynamics of DNA with 2-aminopurine: Effects of local environment on fluorescence, *Biochemistry* **40**, 946–956.
- Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R., and Ross, J. B. (2001) Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine, *Biochemistry* **40**, 957–967.
- Jiang, Y. L., Song, F., and Stivers, J. T. (2002) Base flipping mutations of uracil DNA glycosylase: Substrate rescue using a pyrene nucleotide wedge, *Biochemistry* **41**, 11248–11254.
- Gowher, H., and Jeltsch, A. (2000) Molecular enzymology of the ecorv DNA-(adenine-n (6))-methyltransferase: Kinetics of DNA binding and bending, kinetic mechanism and linear diffusion of the enzyme on DNA, *J. Mol. Biol.* **303**, 93–110.
- Su, T. J., Connolly, B. A., Darlington, C., Mallin, R., and Dryden, D. T. (2004) Unusual 2-aminopurine fluorescence from a complex of DNA and the ecorv methyltransferase, *Nucleic Acids Res.* **32**, 2223–2230.
- Wong, I., Lundquist, A. J., Bernards, A. S., and Mosbaugh, D. W. (2002) Presteady-state analysis of a single catalytic turnover by *Escherichia coli* uracil-DNA glycosylase reveals a “pinch-pull-push” mechanism, *J. Biol. Chem.* **277**, 20.
- Vallur, A. C., Feller, J. A., Abner, C. W., Tran, R. K., and Bloom, L. B. (2002) Effects of hydrogen bonding within a damaged base pair on the activity of wild type and DNA-intercalating mutants of human alkyladenine DNA glycosylase, *J. Biol. Chem.* **277**, 31673–31678.
- Liu, P., Burdzy, A., and Sowers, L. C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: Ung, mug, and tdg, *Chem. Res. Toxicol.* **15**, 1001–1009.
- Valinluck, V., Liu, P., Burdzy, A., Ryu, J., and Sowers, L. C. (2002) Influence of local duplex stability and n(6)-methyladenine on uracil recognition by mismatch-specific uracil-DNA glycosylase (mug), *Chem. Res. Toxicol.* **15**, 1595–1601.
- Biswas, T., Clos, L. J., II, SantaLucia, J., Jr., Mitra, S., and Roy, R. (2002) Binding of specific DNA base-pair mismatches by n-methylpurine-DNA glycosylase and its implication in initial damage recognition, *J. Mol. Biol.* **320**, 503–513.
- Chepanoske, C. L., Langelier, C. R., Chmiel, N. H., and David, S. S. (2000) Recognition of the nonpolar base 4-methylindole in DNA by the DNA repair adenine glycosylase muty, *Org. Lett.* **2**, 1341–1344.
- O’Brien, P. J., and Ellenberger, T. (2004) The *Escherichia coli* 3-methyladenine DNA glycosylase alkA has a remarkably versatile active site, *J. Biol. Chem.* **279**, 26876–26884.
- O’Brien, P. J., and Ellenberger, T. (2004) Dissecting the broad substrate specificity of human 3-methyladenine-DNA glycosylase, *J. Biol. Chem.* **279**, 9750–9757.
- Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E., and Tainer, J. A. (2004) DNA binding and nucleotide flipping by the human DNA repair protein agt, *Nat. Struct. Mol. Biol.* **11**, 714–720.
- Moran, S., Ren, R. X. F., Sheils, C. J., Rumney, S., and Kool, E. T. (1996) Non-hydrogen bonding “terminator” nucleosides in-

- crease the 3'-end homogeneity of enzymatic rna and DNA synthesis, *Nucleic Acids Res.* 24, 2044–2052.
30. Ren, R. X. F., Chaudhuri, N. C., Paris, P. L., Rumney, S., and Kool, E. T. (1996) Naphthalene, phenanthrene, and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA, *J. Am. Chem. Soc.* 118, 7671–7678.
31. Fasman, G. D. (1975) *Handbook of biochemistry and molecular biology: Nucleic acids*, Vol. 1, 3rd ed., CRC Press, Boca Raton, FL.
32. Drohat, A. C., Jagadeesh, J., Ferguson, E., and Stivers, J. T. (1999) Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase, *Biochemistry* 38, 11866–11875.
33. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) Free-energy increments for hydrogen-bonds in nucleic-acid base-pairs, *J. Am. Chem. Soc.* 109, 3783–3785.
34. Jiang, Y. L., Kwon, K., and Stivers, J. T. (2001) Turning on uracil-DNA glycosylase using a pyrene nucleotide switch, *J. Biol. Chem.* 276, 42347–42354.
35. Jiang, Y. L., McDowell, L. M., Poliks, B., Studelska, D. R., Cao, C., Potter, G. S., Schaefer, J., Song, F., and Stivers, J. T. (2004) Recognition of an unnatural difluorophenyl nucleotide by uracil DNA glycosylase, *Biochemistry* 43, 15429–15438.
36. Matray, T. J., and Kool, E. T. (1999) A specific partner for abasic damage in DNA, *Nature* 399, 704–708.
37. Bianchet, M. A., Seiple, L. A., Jiang, Y. L., Ichikawa, Y., Amzel, L. M., and Stivers, J. T. (2003) Electrostatic guidance of glycosylation migration along the reaction coordinate of uracil DNA glycosylase, *Biochemistry* 42, 12455–12460.
38. Werner, R. M., Jiang, Y. L., Gordley, R. G., Jagadeesh, G. J., Ladner, J. E., Xiao, G., Tordova, M., Gilliland, G. L., and Stivers, J. T. (2000) Stressing-out DNA? The contribution of serine-phosphodiester interactions in catalysis by uracil DNA glycosylase, *Biochemistry* 39, 12585–12594.
39. Kwon, K., Jiang, Y., and Stivers, J. (2003) Rational engineering of a DNA glycosylase specific for an unnatural cytosine:pyrene base pair, *Chem. Biol.* 10, 1–20.
40. Moe, J. G., and Russu, I. M. (1992) Kinetics and energetics of base-pair opening in 5'-d(CGCGAATTCGCG)-3' and a substituted dodecamer containing g.T mismatches, *Biochemistry* 31, 8421–8428.
41. Moe, J. G., and Russu, I. M. (1990) Proton exchange and base pair opening kinetics in 5'-d(CGCGAATTCGCG)-3' and related dodecamers, *Nucleic Acids Res.* 18, 821–827.
42. Gueron, M., Charretier, E., Hagerhorst, J., Kochoyan, M., Leroy, J. L., and Moraillon, A. (1990) in *Structural methods* (Sarma, R. H., Ed.) pp 113–137, Adenine Press, New York.
43. Gueron, M., and Leroy, J.-L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange, *Methods Enzymol.* 261, 383–413.
44. Dong, J., Drohat, A. C., Stivers, J. T., Pankiewicz, K. W., and Carey, P. R. (2000) Raman spectroscopy of uracil DNA glycosylase-DNA complexes: Insights into DNA damage recognition and catalysis, *Biochemistry* 39, 13241.
45. Drohat, A. C., and Stivers, J. T. (2000) NMR evidence for an unusually low n1 pKa for uracil bound to uracil DNA glycosylase: Implications for catalysis, *J. Am. Chem. Soc.* 122, 1840–1841.
46. Drohat, A. C., Xiao, G., Tordova, M., Jagadeesh, J., Pankiewicz, K. W., Watanabe, K. A., Gilliland, G. L., and Stivers, J. T. (1999) Heteronuclear NMR and crystallographic studies of wild-type and h187q *Escherichia coli* uracil DNA glycosylase: Electrophilic catalysis of uracil expulsion by a neutral histidine 187, *Biochemistry* 38, 11876–11886.
47. Ikuta, S., Eritja, R., Kaplan, B. E., and Itakura, K. (1987) NMR studies of the stable mismatch purine-thymine in the self-complementary d(CGpuaattcg) duplex in solution, *Biochemistry* 26, 5646–5650.
48. Clore, G. M., Oschkinat, H., McLaughlin, L. W., Benseler, F., Happ, C. S., Happ, E., and Gronenborn, A. M. (1988) Refinement of the solution structure of the DNA dodecamer 5'-d(CGCGpattcgcg)2 containing a stable purine-thymine base pair: Combined use of nuclear magnetic resonance and restrained molecular dynamics, *Biochemistry* 27, 4185–4197.
49. Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) Free-energy contributions of g.U and other terminal mismatches to helix stability, *Biochemistry* 25, 3209–3213.
50. Guckian, K. M., Schweitzer, B. A., Ren, R. X. F., Sheils, C. J., Tahmassebi, D. C., and Kool, E. T. (2000) Factors contributing to aromatic stacking in water: Evaluation in the context of DNA, *J. Am. Chem. Soc.* 122, 2213–2222.
51. Clegg, R. M., Murchie, A. I., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D. M. (1992) Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction, *Biochemistry* 31, 4846–4856.
52. Hiller, D. A., Fogg, J. M., Martin, A. M., Beechem, J. M., Reich, N. O., and Perona, J. J. (2003) Simultaneous DNA binding and bending by ecorv endonuclease observed by real-time fluorescence, *Biochemistry* 42, 14375–14385.
53. Lorenz, M., Hillisch, A., Goodman, S. D., and Diekmann, S. (1999) Global structure similarities of intact and nicked DNA complexed with ihf measured in solution by fluorescence resonance energy transfer, *Nucleic Acids Res.* 27, 4619–4625.
54. Vamosi, G., Gohlke, C., and Clegg, R. M. (1996) Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: Multiple conformers of single-stranded and double-stranded dye-DNA complexes, *Biophys. J.* 71, 972–994.
55. Hopkins, B. B., and Reich, N. O. (2004) Simultaneous DNA binding, bending, and base flipping: Evidence for a novel m.Ecori methyltransferase-DNA complex, *J. Biol. Chem.* 279, 37049–37060.
56. Huang, N., and MacKerell, A. D., Jr. (2005) Specificity in protein-DNA interactions: Energetic recognition by the (cytosine-c5)-methyltransferase from hhai, *J. Mol. Biol.* 345, 265–274.
57. Cheong, C., Tinoco, I., Jr., and Chollet, A. (1988) Thermodynamic studies of base pairing involving 2,6-diaminopurine, *Nucleic Acids Res.* 16, 5115–5122.
58. Chazin, W. J., Rance, M., Chollet, A., and Leupin, W. (1991) Comparative NMR analysis of the decadeoxynucleotide d-(gcattaatgc)2 and an analogue containing 2-aminoadenine, *Nucleic Acids Res.* 19, 5507–5513.
59. Carbonnaux, C., Fazakerley, G. V., and Sowers, L. C. (1990) An NMR structural study of deaminated base pairs in DNA, *Nucleic Acids Res.* 18, 4075–4081.

BI050084U

Uracil-Directed Ligand Tethering: An Efficient Strategy for Uracil DNA Glycosylase (UNG) Inhibitor Development

Yu Lin Jiang, Daniel J. Krosky, Lauren Seiple, and James T. Stivers*

Contribution from the Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received August 25, 2005; E-mail: jstivers@jhmi.edu

Abstract: Uracil DNA glycosylase (UNG) is an important DNA repair enzyme that recognizes and excises uracil bases in DNA using an extrahelical recognition mechanism. It is emerging as a desirable target for small-molecule inhibitors given its key role in a wide range of biological processes including the generation of antibody diversity, DNA replication in a number of viruses, and the formation of DNA strand breaks during anticancer drug therapy. To accelerate the discovery of inhibitors of UNG we have developed a uracil-directed ligand tethering strategy. In this efficient approach, a uracil aldehyde ligand is tethered via alkyloxyamine linker chemistry to a diverse array of aldehyde binding elements. Thus, the mechanism of extrahelical recognition of the uracil ligand is exploited to target the UNG active site, and alkyloxyamine linker tethering is used to randomly explore peripheral binding pockets. Since no compound purification is required, this approach rapidly identified the first small-molecule inhibitors of human UNG with micromolar to submicromolar binding affinities. In a surprising result, these uracil-based ligands are found not only to bind to the active site but also to bind to a second uncompetitive site. The weaker uncompetitive site suggests the existence of a transient binding site for uracil during the multistep extrahelical recognition mechanism. This very general inhibitor design strategy can be easily adapted to target other enzymes that recognize nucleobases, including other DNA repair enzymes that recognize other types of extrahelical DNA bases.

Introduction

DNA repair pathways have been traditionally viewed as the cellular quality control machinery that preserves the coding potential of genomes.¹ However, there is emerging recognition that the repair mechanisms evolved to prevent accumulation of the RNA base uracil in DNA play a much broader role in a number of important areas of biomedicine that are divergent from genome preservation. Remarkable examples include the role of the uracil excision repair machinery in the process of generating genetic diversity during antibody maturation in B cells,^{2–4} the importance of uracil incorporation and removal in the life cycles of herpes,⁵ cytomegalo,⁶ pox,^{7,8} and type 1 human immunodeficiency viruses (HIV-1),⁹ and the essential role of this pathway in generating pharmacologically active single and double strand DNA breaks during chemotherapy treatment with 5-fluorouracil and methotrexate.^{10,11} The key enzyme player in

all of these remarkably diverse processes is uracil DNA glycosylase (UNG), which cleaves the glycosidic bond between the uracil base and the deoxyribose sugar in DNA by flipping the uracil nucleotide from the DNA duplex into the enzyme active site (Figure 1A).¹² Given that UNG is emerging as a very interesting pharmacologic target, we have sought out methods for the rapid and efficient identification of small-molecule ligands that could inhibit its activity. Although potent nucleic acid-based and proteinaceous inhibitors are available that target UNG,^{13–17} there are no small-molecule inhibitors for this enzyme, and strategies for the discovery of such ligands are lacking.

One of the most exciting potential applications of small-molecule human UNG inhibitors are as antiretroviral agents. Recent findings have established that HIV-1 specifically packages human UNG (hUNG) into virus particles via interaction with the virus encoded integrase protein (Int) or perhaps a ternary complex between UNG, Int, and the viral Vpr

- (1) Lindahl, T.; Wood, R. D. *Science* **1999**, *286*, 1897–1905.
- (2) Di Noia, J.; Neuberger, M. S. *Nature* **2002**, *419*, 43–48.
- (3) Imai, K.; Slupphaug, G.; Lee, W. I.; Revy, P.; Nonoyama, S.; Catalan, N.; Yel, L.; Forveille, M.; Kavli, B.; Krokan, H. E.; Ochs, H. D.; Fischer, A.; Durandy, A. *Nat. Immunol.* **2003**, *4*, 1023–1028.
- (4) Storb, U.; Stavnezer, J. *Curr. Biol.* **2002**, *12*, R725–727.
- (5) Chen, R.; Wang, H.; Mansky, L. M. *J. Gen. Virol.* **2002**, *83*, 2339–2345.
- (6) Prichard, M. N.; Duke, G. M.; Mocarski, E. S. *J. Virol.* **1996**, *70*, 3018–3025.
- (7) De Silva, F. S.; Moss, B. J. *Virol.* **2003**, *77*, 159–166.
- (8) Stuart, D. T.; Upton, C.; Higman, M. A.; Niles, E. G.; McFadden, G. J. *Virol.* **1993**, *67*, 2503–2512.
- (9) Priet, S.; Gros, N.; Navarro, J. M.; Boretto, J.; Canard, B.; Querat, G.; Sire, J. *Mol. Cell* **2005**, *17*, 479–490.
- (10) Ladner, R. D. *Curr. Protein Pept. Sci.* **2001**, *2*, 361–370.

- (11) Tinkelenberg, B. A.; Hansbury, M. J.; Ladner, R. D. *Cancer Res.* **2002**, *62*, 4909–4915.
- (12) Stivers, J. T.; Drohat, A. C. *Arch. Biochem. Biophys.* **2001**, *396*, 1–9.
- (13) Bianchet, M. A.; Seiple, L. A.; Jiang, Y. L.; Ichikawa, Y.; Amzel, L. M.; Stivers, J. T. *Biochemistry* **2003**, *42*, 12455–12460.
- (14) Jiang, Y. L.; Cao, C.; Stivers, J. T.; Song, F.; Ichikawa, Y. *Bioorg. Chem.* **2004**, *32*, 244–262.
- (15) Krosky, D. J.; Song, F.; Stivers, J. T. *Biochemistry* **2005**, *44*, 5949–5959.
- (16) Sekino, Y.; Bruner, S. D.; Verdine, G. L. *J. Biol. Chem.* **2000**, *275*, 36506–36508.
- (17) Putnam, C. D.; Shroyer, M. J. N.; Lundquist, A. J.; Mol, C. D.; Arvai, A. S.; Mosbaugh, D. W.; Tainer, J. A. *J. Mol. Biol.* **1999**, *287*, 331–346.

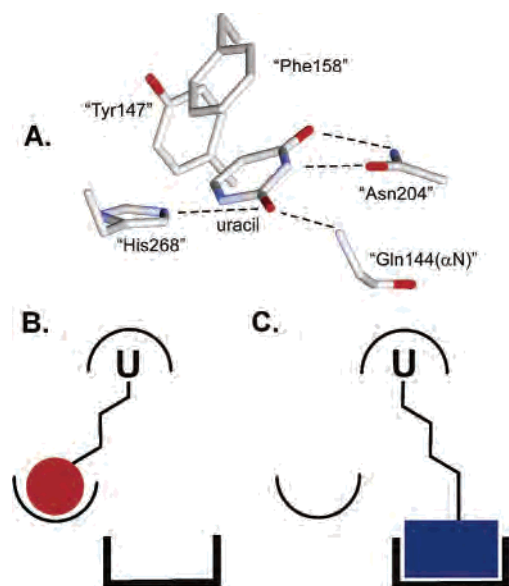


Figure 1. Extrahelical binding of uracil to the UNG active site and the general strategy for uracil-directed ligand tethering. (A) Structure of UNG bound to uracil is shown (pdb code 2eug). The residue numbering is for the human enzyme. (B, C) The uracil ligand (U) that targets the UNG active site is covalently tethered to two different ligands that can interact with distinct binding surfaces near the active site.

protein.^{5,18–25} hUNG is required for infection of nondividing cells such as macrophages and resting T cells and helps maintain a viral reservoir in the host that is crucial for virus spread to the lymphoid organs and T-helper lymphocytes and, ultimately, AIDS pathogenesis.^{20,26} UNG is apparently recruited to minimize uracil incorporation into the viral genome in these cells, which have naturally high levels of dUTP, a good substrate for the viral reverse transcriptase.²⁷ In the absence of UNG, the HIV-1 mutation rate is found to increase by 18-fold resulting in extremely inefficient virus replication in nondividing cells,²⁰ and the virus particles produced from UNG-depleted cells are incapable of infecting new target cells.^{9,28} Pharmacologic targeting of a human enzyme required for virus infectivity is extremely attractive because such a target would not be susceptible to the same high mutagenesis rate and resulting drug resistance as viral encoded proteins.²⁹ Targeting the human enzyme is a viable therapeutic strategy because it is not an essential enzyme. Thus, UNG knock-out mice display no remarkable phenotype, nor do UNG null yeast or human cell lines.³⁰

Herein, we report an integrated high-throughput (HTP) platform for discovering small-molecule ligands that inhibit UNG. The strategy takes advantage of the extrahelical uracil recognition mechanism of UNG by using the specificity and binding energy of a uracil ligand to target the UNG active site^{14,31,32} and then covalent tethering of random functional groups for exploration of nearby binding pockets (Figure 1B). Library members can be rapidly screened using a robust HTP activity assay, and initial hits are quickly optimized using subsequent structure–activity studies. This tethering approach, which uses efficient oxime chemistry (Figure 2), is related to the “combinatorial target-guided ligand assembly” method of Ellman et al.³³ but differs in that the uracil ligand specifically targets the active site rather than irrelevant regions of the enzyme. Thus, the hit-rate and binding affinities of early hits are higher than the more random approach of Ellman and colleagues. This synthetic and screening strategy should be easily adaptable for the discovery of inhibitors of other enzymes that recognize extrahelical bases in DNA or free nucleosides.

Results and Discussion

Synthesis of Uracil-Tethered Oxime Libraries and General Strategy. We sought an inhibitor development strategy that allowed rapid and economical synthesis of small-molecule ligands that explore binding sites near the UNG active site and which could be used directly in HTP screening applications without purification. One efficient synthesis strategy that meets these criteria is outlined in Figure 2. First, flexible diaminoalkanediol linkers of variable length are synthesized from the corresponding dibromoalkanes (Figure 2A). Then the linkers are used to tether uracil aldehyde binding elements (**1–3**) to a library of aldehyde binding elements (RCHO) via the formation of stable oxime linkages (Figure 2B). Each tethering reaction is carried out in one well of a 96-well microtiter plate that contains one equivalent uracil aldehyde, one equivalent RCHO library member, and a mixture of diaminoalkanediol linkers ($n = 2–6$). The reactions typically proceed to 85–99% completion after overnight incubation (DMSO solvent, 37 °C) and produce a 1:2:1 statistical mixture of the homodimeric (U∧U, R∧R) and heterodimeric (U∧R) oximes for each of the five linker lengths present (see Experimental Section and Supporting Information Figure S1). Although two geometric configurations are possible, oxime derivatives with bulky substituents are generally found to be ≥95% in the trans configuration.³⁴ The unpurified oxime mixtures were directly screened for inhibition of UNG at ~100 μM total oxime concentration to ensure that each component in the mixture is present at a concentration in the range 5–10 μM. If significant inhibition is observed by any mixture, the linker length and RCHO binding element that gave rise to the inhibition can be identified by resynthesis of the individual oximes using a single linker length in each reaction (see below).

An important aspect of this approach is that the uracil homodimers present in some reaction mixtures are inhibitory even in the absence of any active heterodimer. For instance,

- (18) Bouhamdan, M.; Benichou, S.; Rey, F.; Navarro, J. M.; Agostini, I.; Spire, B.; Camonis, J.; Slupphaug, G.; Vigne, R.; Benarous, R.; Sire, J. *J. Virol.* **1996**, *70*, 697–704.
- (19) BouHamdan, M.; Xue, Y.; Baudat, Y.; Hu, B.; Sire, J.; Pomerantz, R. J.; Duan, L. X. *J. Biol. Chem.* **1998**, *273*, 8009–8016.
- (20) Chen, R.; Le Rouzic, E.; Kearney, J. A.; Mansky, L. M.; Benichou, S. *J. Biol. Chem.* **2004**, *279*, 28419–28425.
- (21) Klarmann, G. J.; Chen, X.; North, T. W.; Preston, B. D. *J. Biol. Chem.* **2003**, *278*, 7902–7909.
- (22) Mansky, L. M.; Preveral, S.; Selig, L.; Benarous, R.; Benichou, S. *J. Virol.* **2000**, *74*, 7039–7047.
- (23) Payne, S. L.; Elder, J. H. *Curr. Protein Pept. Sci.* **2001**, *2*, 381–388.
- (24) Selig, L.; Benichou, S.; Rogel, M. E.; Wu, L. I.; Vodicka, M. A.; Sire, J.; Benarous, R.; Emerman, M. *J. Virol.* **1997**, *71*, 4842–4846.
- (25) Willetts, K. E.; Rey, F.; Agostini, I.; Navarro, J. M.; Baudat, Y.; Vigne, R.; Sire, J. *J. Virol.* **1999**, *73*, 1682–1688.
- (26) Mansky, L. M.; Le Rouzic, E.; Benichou, S.; Gajary, L. C. *J. Virol.* **2003**, *77*, 2071–2080.
- (27) Miller, R. J.; Cairns, J. S.; Bridges, S.; Sarver, N. *J. Virol.* **2000**, *74*, 7187–7195.
- (28) Elder, R. T.; Zhu, X.; Priet, S.; Chen, M.; Yu, M.; Navarro, J. M.; Sire, J.; Zhao, Y. *Biochem. Biophys. Res. Commun.* **2003**, *306*, 693–700.
- (29) Lau, A.; Swinbank, K. M.; Ahmed, P. S.; Taylor, D. L.; Jackson, S. P.; Smith, G. C.; O'Connor, M. *J. Nat. Cell Biol.* **2005**, *7*, 493–500.

- (30) Nilsen, H.; Rosewell, I.; Robins, P.; Skjelbred, C. F.; Andersen, S.; Slupphaug, G.; Daly, G.; Krokan, H. E.; Lindahl, T.; Barnes, D. E. *Mol. Cell* **2000**, *5*, 1059–1065.
- (31) Drohat, A. C.; Stivers, J. T. *Biochemistry* **2000**, *39*, 11865–11875.
- (32) Drohat, A. C.; Stivers, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 1840–1841.
- (33) Maly, D. J.; Choong, I. C.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2419–2424.
- (34) Winans, K. A.; Bertozzi, C. R. *Chem. Biol.* **2002**, *9*, 113–129.

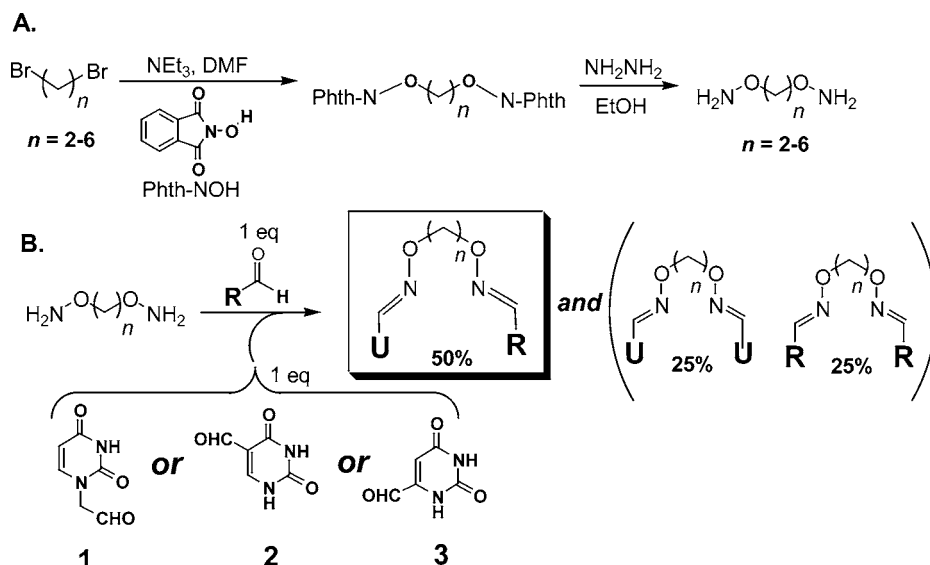


Figure 2. Synthesis of oxime libraries based on uracil and RCHO: (A) synthesis of diaminoalkanediol tethers of variable length; (B) construction of the uracil–oxime library based on the uracil aldehydes (**1–3**) and a series of aldehyde compounds (RCHO). The products consist of a 1:2:1 mixture of the heterodimer (U∧R) and the two homodimers (U∧U and R∧R) connected via alkane linkers of lengths 2–6. A 1 equiv amount of total diaminoalkanediol is added to each reaction. Each linker length is present at one-fifth of the total concentration.

the purified homodimers of various lengths that are based on 6-formyluracil (**3**) give rise to about 22% inhibition in all the mixtures based on **3** under the screening conditions (not shown). In contrast, the homodimers of **1** and **2** show no detectable inhibition under the same conditions. Thus, the screening assay must be robust enough to detect any *additional* inhibition resulting from an active heterodimer in the mixture. Spectroscopic results for determining the purity and composition of representative reaction mixtures are available (see Supporting Information).

High-Throughput Screening of Uracil–Oxime Libraries.

To test this directed library approach, we tethered the three uracil aldehydes (**1–3**) shown in Figure 2 to 14 aldehyde binding elements (RCHO) using the variable-length diaminoalkanediol linkers (see Supporting Information Table S1 for RCHO structures). This library of uracil-linked binding elements was screened for inhibition of hUNG using a high-throughput molecular beacon activity assay (Figure 3).³⁵ In this assay, one DNA strand is labeled with a fluorescent 5'-FAM and the complementary strand is modified with a 3'-dabsyl moiety that serves to efficiently quench the fluorescence of the FAM group through contact quenching. To increase stability, the two DNA strands are linked in a hairpin configuration using an 18 atom poly(ethylene glycol) linker. When the substrate DNA is exposed to UDG, multiple uracils are removed, and eventually the two paired strands of the hairpin spontaneously separate, thus removing the dabsyl quencher from the proximity of the FAM group and resulting in a 6-fold increase in the fluorescence of the system (Figure 3A). Under the assay conditions, the hairpin DNA substrate has a $K_m = 164 \pm 10$ nM and $k_{\text{cat}} = 0.33 \pm 0.01$ s⁻¹ (Figure 3B). To enhance detection of competitive inhibitors during HTP screening we employed a molecular beacon substrate concentration equivalent to $1/3K_m$ (50 nM). Representative HTP screening results for several inactive and active oxime mixtures are shown in Figure 4 ([total oxime] = 100 μM).

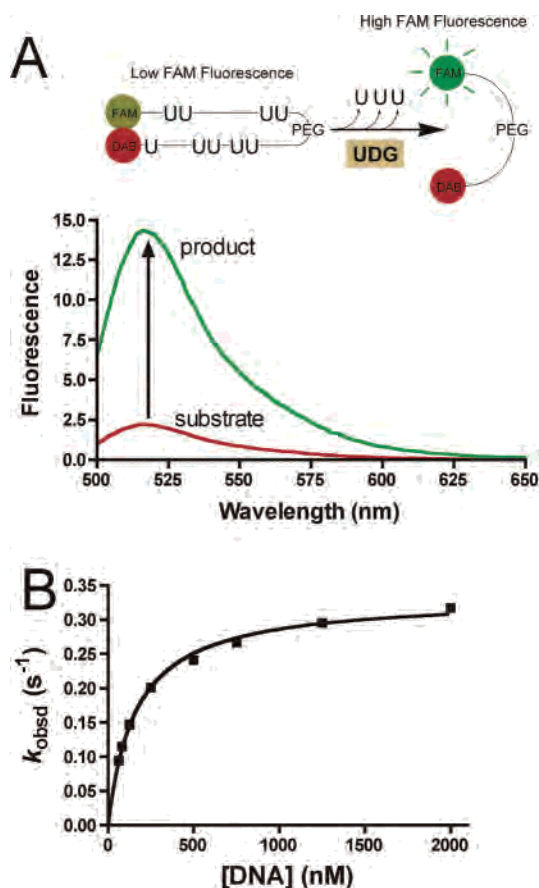


Figure 3. High-throughput (HTP) UDG kinetic assay. (A) The HTP assay relies on molecular beacon technology. Excision of multiple uracil bases by the enzyme destabilizes the hairpin structure thereby releasing the 5'-FAM fluorophore from the quenching effects of the 3'-dabsyl group. (B) Steady-state kinetic analysis is shown of the hUDG reaction using the molecular beacon hairpin substrate.

Several activity trends emerged immediately from the screening results shown in Figure 4. First, none of the mixtures derived from the uracil N1-acetaldehyde binding element (**1**) were inhibitory at the concentration used in the screen. In addition,

(35) Kwon, K.; Nagarajan, R.; Stivers, J. T. *Biochemistry* **2004**, *43*, 14994–15004.

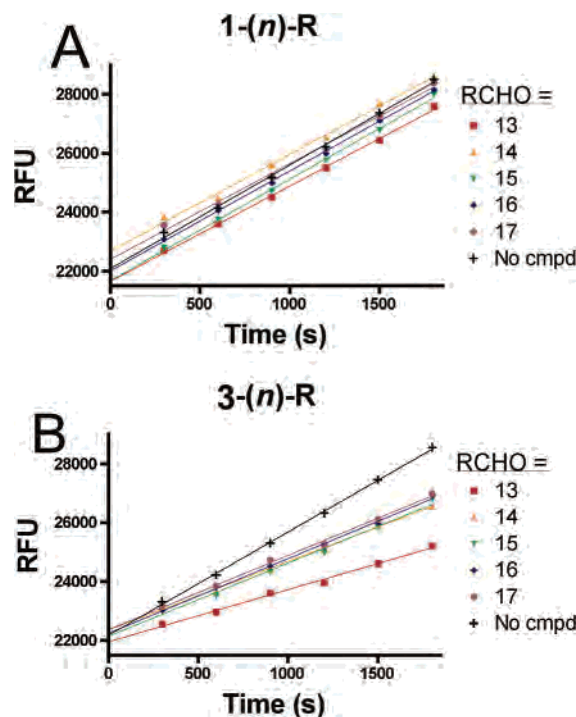


Figure 4. Representative HTP screening results using the molecular beacon substrate. (A) Screen of oxime dimer mixtures derived from uracil aldehyde **1** and aryl aldehydes **13–17** is presented. No inhibition was observed for any oxime derived from **1** regardless of linker length (n). (B) Screen of oxime dimer mixtures derived from uracil aldehyde **3** and aryl aldehydes **13–17** is shown. The mixed oxime derived from **3** and **13** shows significant inhibition, and this derivative was further optimized. For **14–17**, the observed inhibition represents that from the **3–3** homodimers that are present in the mixtures.

none of the U \wedge U homodimers derived from **2** were found to be inhibitory, nor were any of the R \wedge R homodimers regardless of the linker length. (Inhibition by the homodimers is automatically assessed because these are present in multiple reaction mixtures.) In contrast, one oxime mixture derived from uracil aldehydes **2** and **3** and RCHO = 3,4-dihydroxybenzaldehyde (**13**) showed inhibitory activity in the range 15–100%, indicating that active heterodimers were present. The structures of the active heterodimers present in these two oxime mixtures are shown at the top of Table 1.

The two active mixtures were deconvoluted with respect to linker length by individually synthesizing each oxime dimer using a *single* diaminoalkanediol linker/reaction (Table 1). At this stage we did not separate the homodimers from the active heterodimers in the mixtures. For the oxime dimers derived from 5-formyluracil (**2**) and 3,4-dihydroxybenzaldehyde (**13**), a broad dependence on linker length was observed with length $n = 2$ being most favorable for inhibitory activity (i.e. mixed oxime **2-(2)-13**, Chart 1). In contrast, a very stringent linker length of $n = 3$ was required for maximal inhibitory activity with the oxime mixture derived from 6-formyluracil (**3**) and 3,4-dihydroxybenzaldehyde (**13**) to form mixed-oxime **3-(3)-13** (Chart 1). To confirm these results, **2-(2)-13** and **3-(3)-13** were separated from their respective homodimers using reversed phase HPLC (see Methods), and the concentration dependence of inhibition was determined. The measured IC_{50} values for **2-(2)-13** and **3-(3)-13** were 5.8 and 1.1 μ M, respectively (Figure 5).

Structure–Activity Relationships. In an effort to find more potent inhibitors based on the **3-(3)-13** scaffold, 25 commercially

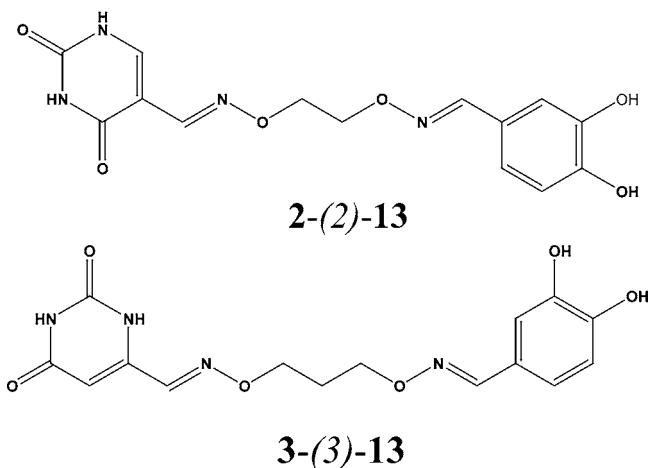
Table 1. Structures of Active Heterodimers and Dependence of Inhibition on Linker Length^a

The chemical structures show two heterodimers. The first, labeled 2-(n)-13, consists of a uracil ring connected via an oxime group to a linker of length n, which is then connected via another oxime group to a 3,4-dihydroxybenzaldehyde ring. The second, labeled 3-(n)-13, consists of a pyrimidine ring connected via an oxime group to a linker of length n, which is then connected via another oxime group to a 3,4-dihydroxybenzaldehyde ring.

mixture	linker length (n)	% inhibition
2-(n)-13	2	50
	3	40
	4	20
	5	20
	6	15
3-(n)-13	2	57
	3	100
	4	51
	5	48
	6	48

^a Reactions were performed in the presence of 100 μ M oxime mixture and 50 nM substrate concentration.

Chart 1. Heterodimer Oximes Identified from Deconvolution of an Active Mixture



available benzaldehyde precursors were purchased (**18–42**; cf. Supporting Information Table S2). The HTP screen was then performed on this set of oxime mixtures (**3-(3)-R**) in an identical fashion as described above. This structure–activity study established that the 3- and 4-hydroxyl groups of **3-(3)-13** were essential for activity because alkylation or halogen substitution at these positions had a substantial deleterious effect on inhibitory activity (see Supporting Information Table S2). Thus,

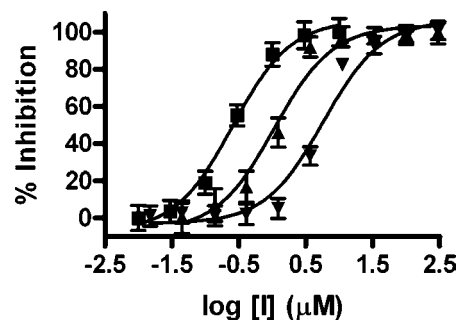


Figure 5. IC_{50} analysis for **2-(2)-13** (\blacktriangledown), **3-(3)-13** (\blacktriangle), and **3-(3)-27** (\blacksquare).

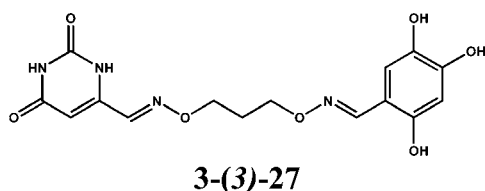
Table 2. Inhibitory Activity for Structural Variants of 3-(3)-27^a

3-(3)-R		
RCHO =		
27	43	44
45	46	
variant	IC ₅₀ (μM)	2-R
3-(3)-27	0.26	OH
3-(3)-43	2.7	F
3-(3)-44	16	Cl
3-(3)-45	40	Br
3-(3)-46	40	NO ₂

^a The concentration dependence of inhibition was determined using 50 nM substrate.

hydrogen bond donating groups at the 3- and 4-positions of the benzyl ring appear to be essential.

One compound in this series with an additional hydroxyl group at the 2-position of the benzyl ring (3-(3)-27) showed a



3-fold greater potency than 3-(3)-13 (Figure 5, ■) (IC₅₀ = 0.3 μM). To further investigate SARs based around the 3-(3)-27 scaffold, we synthesized four more 3,4-dihydroxybenzaldehyde analogues (43–46, Table 2), where the substituent at the 2-position was varied (R = F, Cl, Br, or NO₂). Within this series there was a strong trend correlating with atomic size for the halogens, with the smaller fluorine substituent binding 16-fold more tightly than bromine. However, no substituent in this series was more effective than the 2-hydroxyl group. In conclusion, the binding pocket for the 2-substituent favors a hydrogen bond donating group with a van der Waals radius smaller than chlorine.

Inhibition Mechanisms of 3-(3)-27, 2-(2)-13, and Uracil.

Although the uracil-directed ligand tethering strategy is expected to produce competitive inhibitors of UNG, we thoroughly investigated whether this assumption was true. The detailed mode of inhibition by 3-(3)-27 and 2-(2)-13 was evaluated by varying both substrate and inhibitor concentrations (Figure 6A,B). Standard double reciprocal plots of 1/*k*_{obsd} against 1/[DNA] at increasing concentrations of 3-(3)-27 showed no significant intercept effects establishing a competitive aspect to the inhibition (Figure 6A). However, a secondary plot of the Lineweaver–Burk slopes against [3-(3)-27] showed a *parabolic* response consistent with the presence of at least two inhibitor binding sites (Figure 6A, inset).³⁶ Global discrimination fitting of the inhibition data by computer simulation with the program Dynafit using competitive, noncompetitive, uncompetitive, mixed-type, two-site competitive–noncompetitive, and two-site competitive–uncompetitive inhibition mechanisms unambigu-

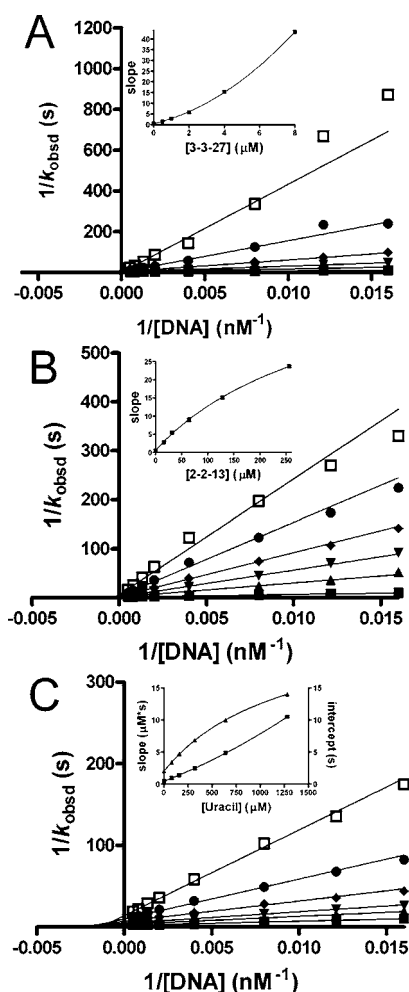


Figure 6. Mode of inhibition analysis, presenting double reciprocal plots and secondary slope and intercept replots for inhibition by increasing concentrations of (A) 3-(3)-27, (B) 2-(2)-13, and (C) uracil. Slope and intercept effects in the inset to (C) are shown as squares and triangles, respectively.

ously confirmed the presence of two inhibitory binding sites for 3-(3)-27 (see Supporting Information).³⁷ Simulations clearly indicated that the first tight site is competitive with respect to substrate. Although the simulations indicated a slight statistical advantage for a partial mixed-type inhibition mode for the second weaker site, it was difficult to eliminate an uncompetitive mode for this site. Using the criterion of Occam's razor, the inhibition parameters for 3-(3)-27 are reported in Table 3 using the simulation results for the competitive–partial uncompetitive mechanism (Scheme 1).

Like its 6-substituted analogue, initial inspection of the Lineweaver–Burk analysis of 2-(2)-13 indicates mixed-type inhibition with a strong preference for binding to the free enzyme (i.e. slope effects, Figure 6B). However, in contrast to 3-(3)-27, the secondary plot of the Lineweaver–Burk slopes versus 2-(2)-13 concentration is *hyperbolic*, indicating that binding of 2-(2)-13 results in partial inhibition (Figure 6B, inset).³⁶ Because binding to the active site would result in complete inhibition, 2-(2)-13 most likely binds to the noncompetitive site observed for 3-(3)-27. Global discrimination fitting of the inhibition data by computer simulation confirmed this inhibition mechanism (Scheme 1) and provided the inhibition

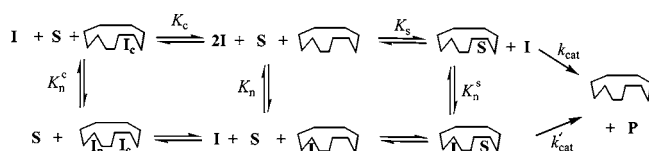
(36) Segel, I. H. *Enzyme Kinetics*; John Wiley & Sons: New York, 1993; Chapter 8, pp 465–504.

(37) Kuzmic, P. *Anal. Biochem.* **1996**, 237, 260–273.

Table 3. Inhibition Constants for Uracil and Its Derivatives^a

param	3-(3)-27	2-(2)-13	uracil	51
K_s (mM)	0.19 ± 0.02	0.23 ± 0.03	0.23 ± 0.02	0.16 ± 0.01
k_{cat} (s^{-1})	0.41 ± 0.01	0.50 ± 0.02	0.47 ± 0.01	0.33 ± 0.01
k_{cat}' (s^{-1})	0.16 ± 0.04	0.012 ± 0.02	0.06 ± 0.01	
K_c (mM)	0.32 ± 0.02		80 ± 7	45 ± 2
K_n (mM)		2.8 ± 0.1		
K_n^c (mM)	1.2 ± 0.2		300 ± 55	
K_n^s (μM)	1 ± 0.3	125 ± 46	104 ± 7	
mode of inhibiti	two sites, competitive, partial uncompetitive	one site, partial mixed-type	two sites, competitive, partial uncompetitive	one site, competitive

^a Parameters correspond to the mechanisms shown in Scheme 1. K_c and K_n represent dissociation constants for inhibitor binding sites that are competitive and noncompetitive with substrate, respectively. K_n^c and K_n^s represent the dissociation constants for inhibitor binding to the noncompetitive site when the active site is occupied by the competitively bound inhibitor or substrate, respectively. In these simulations the Michaelis–Menten parameters for the substrate were fixed using values from nonlinear regression fits (Figure 6). Other parameters were obtained from simulations to the data using the program Dynafit (cf. Supporting Information).

Scheme 1. Inhibition Mechanisms for 3-(3)-27 and 2-(2)-13 and Uracil^a

^a Only 3-(3)-27, 2-(2)-13, and uracil have mechanisms that include the k_{cat}' step. The mechanisms for 3-(3)-27 and uracil do not include the equilibrium constant K_n , and the mechanism for 2-(2)-13 does not include the equilibria K_c or K_n^c .

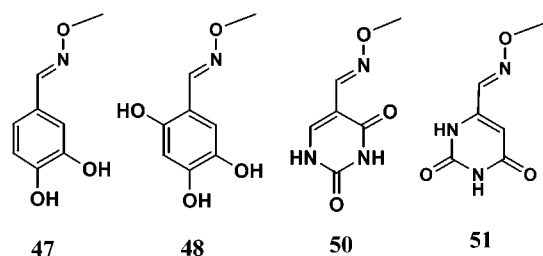
constants reported in Table 3. These observations strongly indicate that 2-(2)-13 binds to a site distinct from the active site, although DNA binding is strongly antagonistic to inhibitor binding (Table 3). In summary, the inhibition mechanisms of 3-(3)-27 and 2-(2)-13 indicate that two inhibition modes exist for these uracil derivatives: one mode competitively targets the active site, and the second weaker mode is noncompetitive or uncompetitive with respect to substrate binding. These data, quite surprisingly, suggested the presence of two uracil binding sites on human UNG.

To further investigate the interesting possibility of two uracil binding sites on UNG, we performed a mode of inhibition analysis for uracil itself (Figure 6C). In confirmation of this initial expectation, inhibition by uracil involves two sites. The first site is competitive, and the second is partially uncompetitive. Accordingly, the Lineweaver–Burk slope replot was slightly parabolic indicating that inhibition involved binding of more than one molecule of uracil, and the intercept replot was hyperbolic indicating a partial uncompetitive mode. These characteristics of the inhibition by uracil combine the features observed for 3-(3)-27 and 2-(2)-13 and establish that the two site binding of 3-(3)-27 is not attributable to the trihydroxybenzaldehyde moiety but, instead, arises from the uracil functionality itself.

Implications for Two Uracil Binding Sites. Why would UNG have a second uracil binding site? Although the answer to this question cannot be firmly established by inhibition data alone, an intriguing role for this site during the mechanism of uracil base flipping is supported by several different experimental findings. First, kinetic experiments following the pathway of uracil flipping from duplex DNA have detected a weakly bound intermediate state of uracil that precedes its attainment of the final extrahelical state seen in the crystal structure (Figure 1A).^{15,38–40} Solution- and solid-state NMR studies of uracil flipping support the existence of a weak uracil binding site

because UNG is found to transiently stabilize thymine and other uracil congeners in an extrahelical conformation, without these bases gaining full access to the uracil active site pocket.^{41,42} Relevant to these observations, the crystal structure of herpesvirus UDG bound to pTTTp shows that the 5' T is bound in the mouth of the active site pocket in a manner that is consistent with a transient state on the pathway for base flipping of uracil.⁴³ Finally, the crystal structure of another base-flipping DNA repair enzyme, human 8-oxoguanine DNA glycosylase, suggests that this related enzyme can flip the normal base guanine into a discrimination pocket that was distinct from the active site pocket that only accommodates 8-oxoguanine.⁴⁴ These combined data provide a compelling case for a generalized pathway for base flipping involving transient enzyme stabilization of at least one extrahelical intermediate state before the base is docked into the active site. On the basis of the observation that 2-(2)-13 is excluded from the active site but that 3-(3)-27 and uracil can occupy both sites, we surmise that the relative binding affinities for each site might depend on the bulkiness of the substituent at the 5-position of uracil. In other words, uracil congeners with small substituents at the five position (such as hydrogen in the case of 3-(3)-27) would favor binding to the active site and uracil derivatives with bulkier substituents (such as the dihydroxybenzaldehyde of 2-(2)-13) would be sterically excluded from the active site but could gain access to the weaker less selective site. Indeed, it is well-known that the active site of UNG uses the bulky side chain of a tyrosine to exclude thymidine (5-methyluracil),^{14,45–47} yet 6-substituted uracil derivatives such as 3-(3)-27 have been generally observed to bind to the active site.¹⁴ Thus, the uracil-based inhibitors found here have revealed a possible pyrimidine discrimination site that may be employed during the multistep extrahelical uracil recognition mechanism. It should be noted that the noncompetitive inhibition

- (38) Jiang, Y. L.; Kwon, K.; Stivers, J. T. *J. Biol. Chem.* **2001**, *276*, 42347–42354.
- (39) Jiang, Y. L.; Song, F.; Stivers, J. T. *Biochemistry* **2002**, *41*, 11248–11254.
- (40) Jiang, Y. L.; Stivers, J. T. *Biochemistry* **2002**, *41*, 11236–11247.
- (41) Cao, C.; Jiang, Y. L.; Stivers, J. T.; Song, F. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1230–1236.
- (42) Jiang, Y. L.; McDowell, L.; Poliks, B.; Studelska, D.; Cao, C.; Potter, G. S.; Schaefer, J.; Song, F.; Stivers, J. T. *Biochemistry* **2004**, *43*, 15429–15438.
- (43) Savva, R.; McAuley-Hecht, K.; Brown, T.; Pearl, L. *Nature* **1995**, *373*, 487–493.
- (44) Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L. *Nature* **2005**, *434*, 612–618.
- (45) Kavli, B.; Slupphaug, G.; Mol, C. D.; Arvai, A. S.; Peterson, S. B.; Tainer, J. A.; Krokan, H. E. *EMBO J.* **1996**, *15*, 3442–3447.
- (46) Mol, C. D.; Arvai, A. S.; Slupphaug, G.; Kavli, B.; Alseth, I.; Krokan, H. E.; Tainer, J. A. *Cell* **1995**, *80*, 869–878.
- (47) Kwon, K.; Jiang, Y.; Stivers, J. *Chem. Biol.* **2003**, *10*, 1–20.

Chart 2. *O*-Methyl Oxime Derivatives of the Aldehyde Binding Elements of **2**-(**2**)-**13** and **3**-(**3**)-**27**

mode for **2**-(**2**)-**13** requires that the final extrahelical state can be attained, albeit inefficiently, even when the transient uracil binding site is occupied by the inhibitor. In contrast, the partial uncompetitive mechanism for binding of **3**-(**3**)-**27** to its second site does not present the same apparent discrepancy, because, for uncompetitive inhibition, the compound binds after the substrate is fully inside the active site pocket (see above).

Inhibition by the Untethered Parts. It is of interest to ask how well uracil-directed ligand tethering has performed. To dissect the energetic contributions of the formyluracil and hydroxybenzaldehyde binding elements of **3**-(**3**)-**27** and **2**-(**2**)-**13**, we synthesized the methyl oxime derivatives of aldehydes **2**, **3**, **27**, and **13** as shown in Chart 2. These methyl oxime derivatives are reasonable mimics of the two individual binding elements and in principle could provide an energetic analysis of the binding affinities of the two separate elements. If the sum of the binding energies of each element equals the entire binding free energy of the whole tethered molecule, then it may be concluded that (i) the tether is energetically inert with respect to binding and (ii) the binding of one element does not affect the other by induced strain or forcing a tighter fit. If the whole tethered molecule binds much more weakly or tightly than expected from the summation of the binding free energies of the two individual binding elements, then nonadditive energetic effects are present. Such effects would indicate either an energetic penalty for tethering (antagonistic binding of the parts) or, alternatively, a nonadditive energetic benefit (synergistic binding of the parts).^{48,49}

Comparison of the binding affinity of **3**-(**3**)-**27** to its competitive site ($K_c^{3-(3)-27} = 0.32 \mu\text{M}$) with that of the 6-formyluracil *O*-methyl oxime binding element alone (**51**) allows estimation of the free energy benefit of tethering the trihydroxybenzaldehyde binding element to the 6-formyluracil oxime part. Conversely, comparison of the binding affinity of **3**-(**3**)-**27** with that of the trihydroxybenzaldehyde *O*-methyl ether (**48**) allows estimation of the free energy benefit of tethering the 6-formyluracil oxime binding element to the trihydroxybenzaldehyde part. The 6-formyluracil *O*-methyl oxime **51** shows a cleanly competitive mode of inhibition with $K_c^{51} = 45 \pm 2 \mu\text{M}$ (Table 4, data not shown). Thus, the enhancement in the free energy of binding upon addition of the trihydroxybenzaldehyde (THB) part to the 6-formyluracil oxime element is $\Delta\Delta G^{\text{THB}} = -RT \ln(K_i^{3-(3)-27}/K_i^{51}) = -3 \text{ kcal/mol}$. We were unable to perform a similar energetic analysis with the trihydroxybenzaldehyde *O*-methyl ether (**48**) due to its extremely weak binding (9%

inhibition at 1 mM concentration, data not shown). Similarly, an energetic analysis of the binding elements comprising **2**-(**2**)-**13** was not possible because of the extremely weak inhibition by the 5-formyluracil *O*-methyl oxime (**50**) and the dihydroxybenzaldehyde *O*-methyl ether (**47**). Nevertheless, the 140-fold greater binding affinity of **3**-(**3**)-**27** as compared to the 6-formyluracil *O*-methyl oxime binding element (**51**) alone indicates that a large benefit can be derived from tethering.⁵⁰

Experimental Section

Reagents and General Methods. All chemicals were purchased from commercial sources without further purification unless otherwise stated. The ^1H , ^{13}C , and ^{19}F NMR spectra were recorded on a 400 MHz Varian Innova instrument. The spectra were recorded in deuteriochloroform (CDCl_3) or in hexadeuteriodimethyl sulfoxide ($\text{DMSO}-d_6$). The chemical shifts of protons are given in ppm with TMS as internal standard. The chemical shifts of carbons are obtained in ppm with solvents as internal standards. That of fluorine is given in ppm with 1% trifluoroacetic acid in $\text{DMSO}-d_6$ as an external standard. Most of oximes were purified by HPLC using aqueous triethylammonium acetate (TEAA) as a running buffer. Therefore, TEAA was not completely removed and it appeared in the NMR spectra. Accordingly, proton and carbon chemical shifts of TEAA were not listed during the characterizations of the oximes. During the purification of the oxime **3**-(**3**)-**27**, 2-mercaptoethanol was used as an antioxidant. Therefore, small amounts of this compound and its oxidation product are also present in the oxime **3**-(**3**)-**27**. Flash chromatographies were performed with silica (70–230 mesh from Sorbent Technologies) and monitored by thin-layer chromatography (TLC) with silica plates (Merck, Kieselgel 60 F254).

Synthesis of Alkyl Hydroxyamines. *O,O'*-Diaminoalkanediol linkers of variable length (ethyl, propyl, butyl, pentyl, hexyl) were prepared from the corresponding dibromoalkanes in two steps according to literature procedures (Figure 2A).^{33,51,52}

General Synthesis of Tethered Oxime Dimers. A set of 14 aryl aldehydes (**4**–**17**; cf. Supporting Information Table S1) was selected for library synthesis for coupling to the three uracil containing aldehydes (**1**–**3**, Figure 2) using the *O,O'*-diaminoalkanediol linkers as follows. To each 0.5-mL well of a Matrix microtiter plate was added a DMSO stock solution of AcOH (20 μL , 150 mM, 3 μmol), uracil aldehyde **1**–**3** (20 μL , 150 mM, 3 μmol), and a single aryl aldehyde (20 μL , 150 mM, 3 μmol). The plate was carefully agitated to make the solutions homogeneous. To each of the uracil–aryl aldehyde mixture was added a DMSO solution of the *O,O'*-diaminoalkanediol linkers containing each of the five linker lengths in equal proportion (22 μL , 150 mM, 3.3 μmol total amine equivalents). The plate was sealed, further agitated, and incubated in an oven for 12 h at 37 $^\circ\text{C}$.

The most potent inhibitors from this first screen **2**-(**2**)-**13** and **3**-(**3**)-**13** were synthesized in larger scale and thoroughly characterized after HPLC purification of the heterodimers as follows.

2-(**2**)-**13**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 8.05 (s, 1 H), 7.91 (s, 1 H), 7.78 (s, 1 H), 7.04 (s, $J = 2.4 \text{ Hz}$, 1 H), 6.86 (m, 1 H), 6.74 (d, $J = 8.0 \text{ Hz}$, 1 H), 4.26 (s, 1 H); ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ 162.40, 151.04, 149.25, 147.92, 145.75, 142.74, 140.66, 123.05, 119.88, 115.74, 113.10, 104.31, 71.82, 71.54; UV/vis λ_{max} 275 nm; HRMS (m/z) [$\text{M} + \text{Na}$] $^+$ calcd for $\text{C}_{14}\text{H}_{14}\text{N}_4\text{O}_6\text{Na}$ 357.08, found 357.08.

3-(**3**)-**13**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.10 (bs, H), 8.01 (s, 1 H), 7.94 (s, 1 H), 7.04 (d, $J = 1.6 \text{ Hz}$, 1 H), 6.82 (d, $J = 7.6 \text{ Hz}$, 1 H), 6.74 (d, $J = 7.6 \text{ Hz}$, 1 H), 5.78 (s, 1 H), 4.26 (t, $J = 6.8 \text{ Hz}$, 2 H), 4.12 (t, $J = 6.0 \text{ Hz}$, 2 H), 2.06 (t, $J = 6.8 \text{ Hz}$, 2 H); ^{13}C NMR (125 MHz,

(48) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1678–1683.

(49) Jencks, W. P. Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect. In *Catalysis in Chemistry and Enzymology*; Dover Publications: New York, 1987; pp 615–807.

(50) The uracil N1-acetaldehyde *O*-methyl oxime (**49**) showed undetectable inhibition ($K_c > 10 \text{ mM}$).

(51) Kung, P. P.; Bharadwaj, R.; Fraser, A. S.; Cook, D. R.; Kawasaki, A. M.; Cook, P. D. *J. Org. Chem.* **1998**, *63*, 1846–1852.

(52) Weiss, R. H.; Furfine, E.; Hausleden, E.; Dixon, D. W. *J. Org. Chem.* **1984**, *49*, 4969–4972.

DMSO- d_6) δ 163.95, 151.15, 148.96, 148.04, 145.89, 144.73, 142.23, 123.12, 119.83, 115.81, 113.15, 101.60, 71.94, 69.76, 28.46; UV/vis λ_{max} 273 nm; HRMS (m/z) [$M + H$] $^+$ calcd for $C_{15}H_{17}N_4O_6$ 349.11, found 349.11.

The second set of oxime dimers based on the **3-(3)-13** hit discovered in the first screening round were synthesized in an identical fashion as described above using uracil aldehyde **3** and hydroxybenzaldehydes **18–42** and the *O,O'*-diaminopropanediol linker (cf. Supporting Information Table S2). The most potent inhibitor identified from this second round of screening (**3-(3)-27**) was synthesized in larger scale and thoroughly characterized. **3-(3)-27**: ^1H NMR (400 MHz, DMSO- d_6) δ 9.10 (bs, H), 8.21 (s, 1 H), 7.94 (s, 1 H), 6.88 (s, 1 H), 6.31 (s, 1 H), 5.78 (s, 1 H), 4.28 (t, $J = 6.0$ Hz, 2 H), 4.10 (t, $J = 6.0$ Hz, 2 H), 2.06 (m, 2 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.91, 151.04, 150.25, 149.17, 146.90, 144.63, 142.20, 138.73, 112.76, 107.78, 103.56, 101.69, 71.90, 69.77, 28.38; UV/vis λ_{max} 286 nm; ESI (m/z) for [$M + H$] $^+$ calcd for $C_{15}H_{18}N_4O_7$ 366, found 366; ESI (m/z) for [$M + Na$] $^+$ calcd for $C_{15}H_{17}N_4O_7Na$ 388, found 388; ESI (m/z) for [$M - H$] $^-$ calcd for $C_{15}H_{16}N_4O_7$ 364, found 364.

Isolation and Purification of Oxime Dimers using HPLC. All of the most active oxime heterodimers were purified by HPLC using a Phenomenex Aqua reversed phase C-18 HPLC column (250 mm, 10 mm, 5 μm). Most of the oximes were purified using gradient elution from 0 to 30% CH_3CN in 0.1 M aqueous TEAA over the course of 2 h using UV detection at 254 nm. An exception was oxime **3-(3)-27**, which is prone to air oxidation. In this case, 25 mM 2-mercaptoethanol was added to both of the running buffers. The oximes all eluted with baseline resolution in the order U–U homodimer, U–R heterodimer, followed by the R–R homodimer. This HPLC method was also used to confirm the expected 1:2:1 stoichiometries of homodimer and heterodimer oxime formation, using 10 representative uracil and aryl aldehydes from the library (see Supporting Information Figure S1). Additional NMR evidence supporting the expected stoichiometries is detailed in the Supporting Information Figures S2 and S3.

Synthesis of 2-R-Substituted 3,4-Dihydroxybenzaldehydes and the Corresponding Mixed Oximes with 3. Aldehyde **43** was synthesized by removing the methyl groups of the commercially available 3,4-dimethoxy-6-fluorobenzaldehyde using BBr_3 in CH_2Cl_2 .⁵³ The aldehydes **44** and **45** were synthesized by removing the methylene group of the corresponding 2-halogenated piperonal using AlCl_3 and 6N HCl .⁵⁴ Aldehyde **46** was commercially available. These four aldehydes (**43–46**) were reacted with 6-formyluracil **3** and the *O,O'*-diaminopropanediol linker using the procedure described above, and the mixed oxime dimer was obtained after HPLC purification.

3-(3)-43: ^1H NMR (400 MHz, DMSO- d_6) δ 8.11 (s, 1 H), 7.94 (s, 1 H), 7.05 (d, $J = 7.2$ Hz, 1 H), 6.58 (d, $J = 6.8$ Hz, 1 H), 5.77 (s, 1 H), 5.10 (bs, H), 4.28 (t, $J = 6.8$ Hz, 2 H), 4.16 (t, $J = 6.0$ Hz, 2 H), 2.07 (m, 2 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.91, 155.50, 153.09, 151.08, 149.87, 149.76, 144.65, 142.72, 142.30, 142.21, 110.91, 110.87, 108.62, 108.49, 103.21, 102.96, 101.67, 71.85, 70.06, 28.38; ^{19}F NMR (DMSO- d_6) δ -54.33, -54.35, -54.36, -54.38; UV/vis λ_{max} 268 nm; HRMS (m/z) [$M + Na$] $^+$ calcd for $C_{15}H_{15}FN_4O_6Na$ 389.09, found 389.09.

3-(3)-44: ^1H NMR (400 MHz, DMSO- d_6) δ 8.23 (s, 1 H), 7.94 (s, 1 H), 7.18 (s, 1 H), 6.77 (s, 1 H), 5.78 (s, 1 H), 4.27 (t, $J = 5.6$ Hz, 2 H), 4.16 (t, $J = 6.0$ Hz, 2 H), 2.08 (m, 2 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.90, 151.08, 149.71, 145.48, 145.16, 144.64, 142.21, 123.00, 119.09, 116.08, 112.33, 101.67, 71.85, 70.22, 28.37; UV/vis λ_{max} 275 nm; HRMS (m/z) [$M + Na$] $^+$ calcd for $C_{15}H_{15}ClN_4O_6Na$ 405.06, found 405.06.

3-(3)-45: ^1H NMR (400 MHz, DMSO- d_6) δ 8.18 (s, 1 H), 7.94 (s, 1 H), 7.19 (s, 1 H), 6.93 (s, 1 H), 5.78 (s, 1 H), 4.27 (t, $J = 6.4$ Hz, 2

H), 4.16 (t, $J = 6.4$ Hz, 2 H), 2.08 (m, 2 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.91, 151.10, 150.07, 147.41, 146.02, 144.66, 142.22, 120.58, 119.15, 112.89, 112.29, 101.69, 71.86, 70.24, 28.38; UV/vis λ_{max} 278 nm; HRMS (m/z) [$M + Na$] $^+$ calcd for $C_{15}H_{15}BrN_4O_6Na$ 449.01, found 449.01.

3-(3)-46: ^1H NMR (400 MHz, DMSO- d_6) δ 8.56 (d, $J = 1.2$ Hz, 1 H), 7.95 (s, 1 H), 7.37 (s, 1 H), 6.74 (s, 1 H), 6.26 (bs, H), 5.78 (d, $J = 1.2$ Hz, 1 H), 4.28 (t, $J = 6.0$ Hz, 2 H), 4.16 (t, $J = 6.0$ Hz, 2 H), 2.09 (m, 2 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.88, 160.27, 151.05, 148.62, 148.00, 144.62, 142.22, 133.36, 122.38, 113.80, 109.23, 101.73, 71.85, 70.07, 28.39; UV/vis λ_{max} 269 nm; HRMS (m/z) [$M + H$] $^+$ calcd for $C_{15}H_{16}N_5O_8$ 394.10, found 394.10.

Synthesis of Methyl Oxime Derivatives of 1–3, 13, and 27. The *O*-methyl oxime of 3,4-dihydroxybenzaldehyde (**47**) is known and was synthesized using **13** and *O*-methylhydroxylamine hydrochloride.⁵⁵ *O*-Methyloximes **48–51** were made using a similar method.

48: To a solution of **27** (308 mg, 2.0 mmol) in 4.0 mL of $\text{EtOH}-\text{H}_2\text{O}-\text{THF}$ (0.45/0.3/0.25) were added sodium acetate (264 mg) and *O*-methylhydroxylamine hydrochloride (183 mg), and the solution was stirred at room temperature for overnight. The solvents were removed in vacuo, and the residue was extracted with chloroform three times. The combined organic layers were dried over anhydrous MgSO_4 and concentrated in vacuo. The residue was purified by silica gel column chromatography ($\text{EtOAc}/\text{hexanes}$) to give a product amount of 347 mg in 95% yield: ^1H NMR (400 MHz, DMSO- d_6) δ 9.38 (s, 1 H), 9.21 (s, 1 H), 8.52 (s, 1 H), 8.18 (s, 1 H), 6.88 (s, 1 H), 6.30 (s, 1 H), 3.80 (s, 3 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 150.17, 148.88, 146.52, 138.61, 112.51, 107.77, 103.46, 61.28; UV/vis λ_{max} 239, 274 nm; HRMS (m/z) [$M + H$] $^+$ calcd for $C_8H_{10}NO_4$ 184.06, found 184.06.

49: To a solution of **1** (10.8 mg, 0.063 mmol) in hot DMF (0.5 mL) were added sodium acetate (5.2 mg, 0.063 mmol) solution in water (0.1 mL) and *O*-methylhydroxylamine hydrochloride (5.3 mg, 0.063 mmol), and the solution was stirred at room temperature for overnight. The solvents were removed in vacuo, and the residue was purified by column chromatography using 10–15% (v/v) methanol in CH_2Cl_2 , resulting in 90% yield (10.3 mg 50/50 mixture of trans and cis geometric isomers): ^1H NMR (400 MHz, chloroform- d) δ 9.56 (s, 1 H), 7.43 (t, $J = 5.2$ Hz, 0.5 H), 7.20 (m, 1 H), 6.80 (t, $J = 4.4$ Hz, 0.5 H), 5.78 (m, 1 H), 4.55 (d, $J = 4.4$ Hz, 1 H), 4.48 (d, $J = 5.6$ Hz, 1 H), 3.94 (s, 1.5 H), 3.87 (s, 1.5 H); ^{13}C NMR (100 MHz, DMSO- d_6) δ 163.92, 163.87, 151.06, 150.99, 144.72, 144.54, 143.93, 143.53, 103.14, 102.97, 62.68, 62.37, 46.52, 43.76; UV/vis λ_{max} 263 nm; HRMS (m/z) [$M + H$] $^+$ calcd for $C_7H_{10}N_3O_3$ 184.07, found 184.07.

50: To a solution of **2** (70 mg, 0.5 mmol) in hot DMF (1 mL) were added sodium acetate (41 mg, 0.5 mmol) solution in water (0.5 mL) and *O*-methylhydroxylamine hydrochloride (42 mg, 0.5 mmol), and the solution was stirred at room temperature for 4 h. The solvents were removed in vacuo, and the residue was collected by filtration and washed with cold water 2×1 mL, resulting in 76% yield (70 mg 87/13 mixture of trans and cis geometric isomers): ^1H NMR (400 MHz, DMSO- d_6) δ 11.40 (bs, 2 H), 8.52 (s, 0.13 H), 7.87 (s, 0.87 H), 7.74 (s, 0.87 H), 7.29 (s, 0.13 H), 3.89 (s, 0.39 H), 3.80 (s, 2.61 H); ^{13}C NMR (100 MHz, DMSO- d_6) δ 162.98, 162.36, 150.80, 150.27, 146.08, 142.26, 140.09, 137.31, 104.41, 103.43, 62.32, 61.44; UV/vis λ_{max} 288 nm; HRMS (m/z) [$M + H$] $^+$ calcd for $C_6H_8N_3O_3$ 170.06, found 170.06.

51: To a solution of **3** (79 mg, 0.5 mmol) in hot DMF (2.0 mL) were added sodium acetate (46 mg, 0.5 mmol) solution in water (0.5 mL) and *O*-methylhydroxylamine hydrochloride (46 mg, 0.5 mmol), and the solution was stirred at 50 $^\circ\text{C}$ for 4 h. The solvents were removed in vacuo, and the residue was washed by cold water. After the filtration, product was obtained in 62% yield (53 mg): ^1H NMR (400 MHz, DMSO- d_6) δ 11.18 (s, 1 H), 10.77 (s, 1 H), 7.91 (s, 1 H), 5.77 (s, 1 H), 3.96 (s, 3 H); ^{13}C NMR (125 MHz, DMSO- d_6) δ 163.87, 151.02,

(53) Kirk, K. L.; Cantacuzene, D.; Nimitkitpaisan, Y.; McCulloh, D.; Padgett, W. L.; Daly, J. W.; Creveling, C. R. *J. Med. Chem.* **1979**, *22*, 1493–1497.

(54) Reitz, A.; Avery, M. A.; Verlander, M. S.; Goodman, M. J. *Org. Chem.* **1981**, *46*, 4859–4863.

(55) Watanabe, T.; Suzuki, T.; Umezawa, Y.; Takeuchi, T.; Otsuka, M.; Umezawa, K. *Tetrahedron* **2000**, *56*, 741–752.

144.51, 142.17, 101.41, 62.84; UV/vis λ_{max} 292 nm; HRMS (m/z) [$M + Na$]⁺ calcd for $C_6H_7N_3O_3Na$ 192.04, found 192.04.

High-Throughput Inhibitor Screening. The substrate in this HTS assay was synthesized using standard phosphoramidite DNA solid-phase chemistry using reagents purchased from Glen Research. The DNA was purified using anion exchange chromatography followed by desalting using reversed phase methods. The sequence and size was confirmed using analytical denaturing polyacrylamide gel electrophoresis and MALDI-MS. The substrate is a double-stranded 14-mer DNA containing nine U•A base pairs (5'-FAM-GCA CUU AAG AAU UG: 3'-DABSYL-CA AUU CUU AAG UGC). The UNG HTS assay is performed as follows. To a 96-well microtiter plate was added 5 μ L (2 mM total) of compound in DMSO, followed by 75 μ L (33.3 pM) of human UNG in reaction buffer (10 mM Tris•HCl, pH 8.0, 20 mM NaCl, 7.5 mM MgCl₂, 0.002% brij-35). The reactions were initiated by the addition of 20 μ L (250 nM) of molecular beacon substrate in reaction buffer. The plates are incubated at ambient temperature in a fluorescence plate reader for 30 min, and the progress of the reaction was monitored every 5 min (ex 485 nm/em 520 nm). The final concentrations of the reagents in the assay are 10 mM Tris•HCl, pH 8.0, 20 mM NaCl, 7.5 mM MgCl₂, 0.002% Brij-35, 25 pM human UNG, 50 nM molecular beacon substrate, 100 μ M total compound, and 5% DMSO. The MgCl₂ is essential to increase the stability of the double-stranded DNA substrate and, thus, decrease the initial fluorescence of the molecular beacon and increase the maximum signal of the assay. Addition of Brij-35, a nonionic detergent, is essential to stabilize human UNG at the low concentration used in this assay. A similar assay has been described by Maksimenko et al. that utilizes a 39-mer hairpin DNA.⁵⁶ However, the synthesis and purification of this more complex substrate proceeds with low efficiency and requires higher temperature to induce strand separation (Krosky and Stivers, unpublished data). In contrast, the 14-mer double-stranded molecular beacon is routine and allows screening to be performed conveniently at room temperature.

Mechanism of Inhibition. The substrate used in mechanism of inhibition studies was a modified DNA hairpin where the two strands described above are connected by a hexakis[poly(ethylene glycol)] linker (PEG-U9). This substrate was easier to synthesize and purify than an all-DNA hairpin and, unlike the double stranded DNA substrate, does not require MgCl₂ to achieve minimum fluorescence. To a 96-well plate was added 5 μ L of compound in DMSO, followed by 75 μ L of PEG-U9 hairpin in reaction buffer (20 mM Tris•HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂, 0.002% Brij-35, 1 mM DTT). Eight different DNA concentrations were used in the range 62.5–2000 nM. Reactions were initiated by the addition of 20 μ L of 0.5 nM human UNG in reaction buffer. The final concentrations of reagents in the assay are 20 mM Tris•HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂, 0.002% Brij-35, 1 mM DTT, 5% DMSO, 0.1 nM human UNG, 62.5–2000 nM

(56) Maksimenko, A.; Ishchenko, A. A.; Sanz, G.; Laval, J.; Elder, R. H.; Saparbaev, M. K., *Biochem. Biophys. Res. Commun.* **2004**, 319, 240–246.

PEG-U9 hairpin DNA, and variable amounts of inhibitor. The plates were incubated at ambient temperature in a fluorescence plate reader for 60 min, and the progress of each reaction was monitored every 5 min (ex 485 nm/em 520 nm). Afterward, *Escherichia coli* UNG was added to each well to drive the reactions to completion, and the overall change in fluorescence values were measured. These values were used to convert initial velocities from units of fluorescence units/s to [product]/s. Mechanisms of inhibition and their corresponding inhibitor dissociation constants were determined by Lineweaver–Burk slope and intercept replot analysis and by computational simulations of the initial velocity against inhibitor concentration data using Dynafit v.3.28 (see Supporting Information)

Conclusions

We have developed an efficient strategy to develop small-molecule inhibitors of UNG that have the potential for activity in cell culture or in vivo. The method is quite general and could be adapted to target other enzymes that bind extrahelical bases or free nucleosides. Two future targets of the current uracil mixed oxime library would be the essential bacterial enzyme deoxyuridine nucleotidylhydrolase, which converts dUTP to dUMP,^{23,57–60} and human thymidine phosphorylase, an enzyme implicated in vascularization of tumors.⁶¹ Such inhibitors could serve as useful tools to study the life cycle of pathogenic human viruses, the biology of uracil base excision repair in normal cell lines and tissues, and mechanisms of tumor vascularization.

Acknowledgment. This work was supported by NIH Grant GM56834-10 to J.T.S. D.J.K was supported by the DOD Breast Cancer Research Program (Grant DAMD17-03-1-1251).

Supporting Information Available: Tables of aryl aldehydes and hydroxybenzaldehydes used to construct oxime libraries, HPLC traces and NMR spectra to ascertain stoichiometries of oxime mixtures, ¹H and ¹³C NMR spectra of **3-(3)-13** and **3-(3)-27**, computer simulations of inhibition data, and analytical equations for each inhibition mechanism. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA055846N

- (57) Grasser, F. A.; Romeike, B. F.; Niedobitek, G.; Nicholls, J.; Kremmer, E. *Curr. Protein Pept. Sci.* **2001**, 2, 349–360.
(58) Hidalgo-Zarco, F.; Gonzalez-Pazanowska, D. *Curr. Protein Pept. Sci.* **2001**, 2, 389–397.
(59) Studebaker, A. W.; Balendiran, G. K.; Williams, M. V. *Curr. Protein Pept. Sci.* **2001**, 2, 371–379.
(60) Williams, M. V. *Virology* **1988**, 166, 262–264.
(61) Toi, M.; Atiqur Rahman, M.; Bando, H.; Chow, L. W. *Lancet Oncol.* **2005**, 6, 158–166.

Synthesis and high-throughput evaluation of triskelion uracil libraries for inhibition of human dUTPase and UNG2

Yu Lin Jiang, Suhman Chung, Daniel J. Krosky and James T. Stivers*

*Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine,
725 North Wolfe Street, Baltimore, MD 21205, USA*

Received 22 February 2006; revised 10 April 2006; accepted 13 April 2006
Available online 5 May 2006

Abstract—Human nuclear uracil DNA glycosylase (UNG2) and deoxyuridine triphosphate nucleotidohydrolase (dUTPase) are the primary enzymes that prevent the incorporation and accumulation of deoxyuridine in genomic DNA. These enzymes are desirable targets for small molecule inhibitors given their roles in a wide range of biological processes ranging from chromosomal rearrangements that lead to cancer, viral DNA replication, and the formation of toxic DNA strand breaks during anticancer drug therapy. To accelerate the discovery of such inhibitors, we have developed a high-throughput approach for directed library synthesis and screening. In this efficient technology, a uracil-aldehyde ligand is covalently tethered to one position of a trivalent alkyloxyamine linker via an oxime linkage, and then the vacant linker positions are derivatized with a library of aldehydes. The resulting triskelion oximes were directly screened for inhibitory activity and the most potent of these showed micromolar binding affinities to UNG2 and dUTPase.

© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

From the classic view of DNA repair and mutagenesis, the uracil base has no place in genomic DNA.¹ Accordingly, elaborate DNA repair mechanisms have evolved to exclude dUTP from the nucleotide pool used for DNA replication,^{5,6} and to remove uracil from DNA when it arises from spontaneous deamination of cytosine bases.⁷ However, the uracil base has recently been found to play a much more diverse role in human biology, disease, and anticancer therapy (Fig. 1). Surprisingly, the uracil excision repair machinery has been found to participate in the process of generating somatic mutations during antibody maturation in B cells,⁸⁻¹⁰ and uracil incorporation and/or removal is critical in the life cycles of herpes,¹¹ cytomegalo,¹² pox,^{13,14} and type 1 human immunodeficiency viruses (HIV-1).¹⁵ Furthermore, this pathway also generates the pharmacologically active single and double strand DNA breaks that are the essential tumor killing lesions produced by the widely used anticancer drugs 5-fluoro-

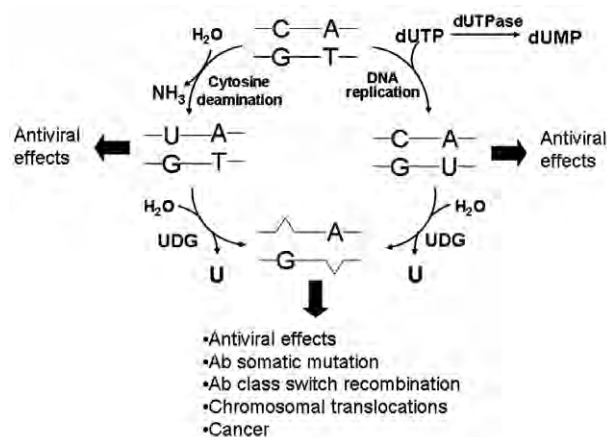


Figure 1. Biological effects of uracil in DNA.²⁻⁴

uracil and methotrexate,^{16,17} and generates the characteristic chromosomal translocations found in some B cell lymphomas. Thus, pharmacologic agents that inhibit these processes are desirable for both investigational and therapeutic purposes.

Human nuclear uracil DNA glycosylase (UNG2) and deoxyuridine triphosphate nucleotidohydrolase

Keywords: Uracil DNA glycosylase; dUTPase; Enzyme inhibition; Directed chemical libraries; High-throughput screening.

*Corresponding author. Tel.: +1 410 502 2758; fax: +1 410 955 3023; e-mail: jstivers@jhmi.edu

(dUTPase) are the primary enzymes that prevent the incorporation and accumulation of deoxyuridine in genomic DNA.^{17,18} Given that these enzymes are emerging as interesting pharmacologic targets, we have sought out methods for the rapid and efficient identification of small molecule ligands that can inhibit their activity. One of the most exciting potential applications of UNG2 and dUTPase inhibitors would be as antiretroviral agents. Recent findings have established that HIV-1 specifically packages UNG2 into virus particles via interaction with the virus encoded integrase protein (Int), or perhaps a ternary complex between UNG, Int and the viral Vpr protein.^{11,19–26} UNG2 is required for infection of nondividing cells such as macrophages and resting T cells, and virus particles produced from UNG depleted cells are incapable of infecting new target cells.^{15,27} Infection of macrophages helps maintain a viral reservoir in the host that is crucial for virus spread to the lymphoid organs and T-helper lymphocytes, and ultimately, AIDS pathogenesis.^{21,28} UNG is apparently recruited to minimize uracil incorporation into the viral genome in these cells, which have naturally high levels of dUTP, a good substrate for the viral reverse transcriptase.²⁹ Inhibition of dUTPase would be expected to further increase dUTP levels in macrophages, resulting in even more uracil misincorporation into the viral genome (Fig. 1). Pharmacologic targeting of a UNG2 and dUTPase is extremely attractive because these targets would not be susceptible to the same high mutagenesis rate and resulting drug resistance as viral encoded proteins.³⁰ Targeting UNG2 is a viable therapeutic strategy because it is not an essential enzyme. Thus, UNG knock-out mice display no remarkable phenotype, nor do UNG null yeast or human cell lines.³¹ Although dUTPase is an essential enzyme in all organisms, it would be expected that rapidly replicating viruses such as HIV-1 would show higher sensitivity than the host, providing a potential therapeutic window.

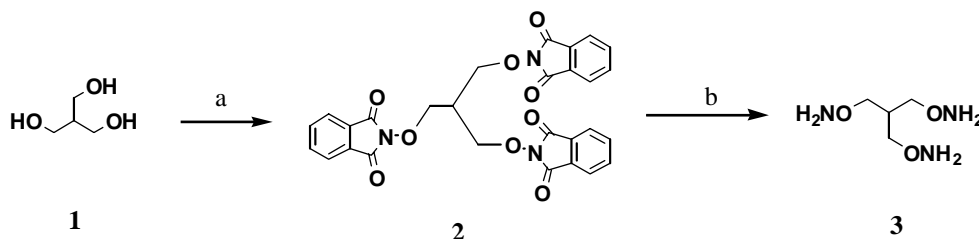
Herein, we report an integrated high-throughput (HTP) platform for the synthesis and evaluation of uracil-directed small molecule libraries based upon triskelion oxyamine scaffolds. The strategy is to attach a uracil-aldehyde ligand to one or two arms of the triskelion scaffold and then derivatize the vacant position(s) with a random library of aldehydes (RCHO). The uracil moiety is expected to weakly target the fully functionalized compound to the active site rather than irrelevant regions of the enzyme, and the random functional groups can then explore nearby binding pockets resulting in increased affinity over that of the uracil alone. Library compounds are rapidly screened using robust HTP activity assays, from which several inhibitors of UNG2 and dUTPase have been identified.

2. Results and discussion

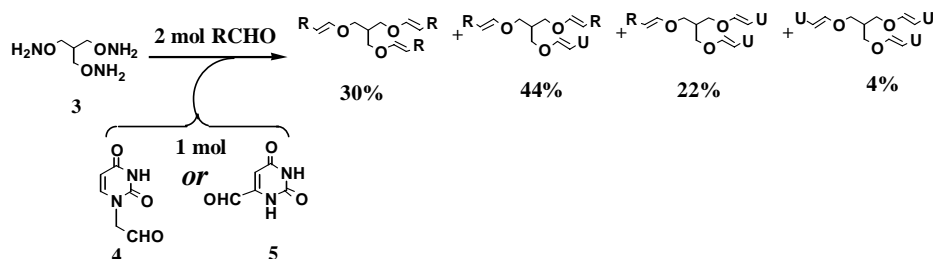
2.1. Synthesis of uracil triskelion oxime libraries

We sought an inhibitor development strategy that allowed rapid and economical synthesis of small molecule ligands that explore binding sites near the UNG2 and dUTPase active sites, and which could be used directly in HTP screening applications without purification. One efficient synthesis strategy that meets these criteria is outlined in Schemes 1 and 2. First, a triskelion oxyamine scaffold is synthesized in two steps from tris(hydroxymethyl)methane (Scheme 1). Then the three oxyamine groups are derivatized with a uracil-aldehyde and a library of 215 aldehyde binding elements (RCHO, Table 1) via the formation of stable oxime linkages (Scheme 2).

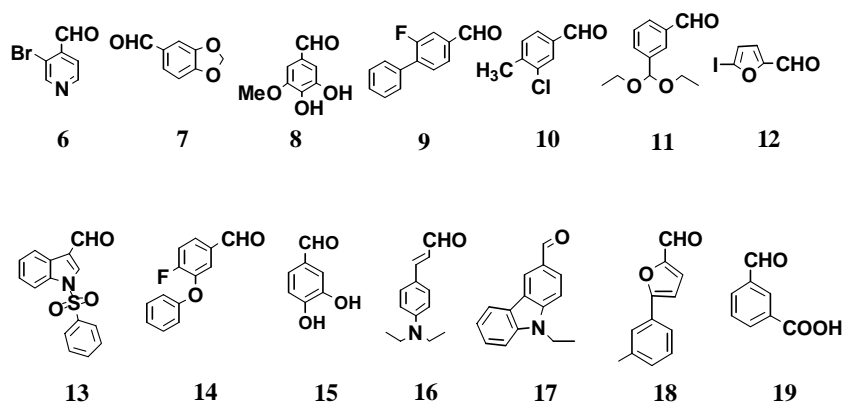
Each linking reaction is carried out in one well of a 96-well microtiter plate that contains one molar equivalent



Scheme 1. Reagents and conditions: (a) triphenylphosphine, diisopropyl azodicarboxylate, *N*-hydroxyphthalimide, anhydrous THF, overnight, 0 °C, 53% yield; (b) anhydrous NH_2NH_2 , 95% ethanol, room temperature, 2 h, 67%. See Ref. 32.



Scheme 2. Reagents and conditions: AcOH (20% v/v), DMSO, 12 h, 37 °C.

Table 1. Representative aryl aldehydes (RCHO) used in library synthesis^a

^a The entire library consisted of 215 aryl aldehydes.

uracil aldehyde, two molar equivalents RCHO library member, and one molar equivalent oxyamine triskelion scaffold (**Scheme 2**). The reactions typically proceed to 85–99% completion after overnight incubation (DMSO, 37 °C), and produce a statistical mixture of the homotrimeric (UUU, RRR) and heterotrimeric (UUR, RRU) oximes in the approximate amounts indicated in **Scheme 2** (see **supplemental materials**). No compound purification is required before screening the library for active compounds.

2.2. High-throughput screening of triskelion libraries against UDG

We have previously developed a high-throughput fluorescence assay for UNG2 that allows screening of chemical libraries.³³ This assay was used to screen the 215 oxime mixtures obtained from reaction of the triskelion oxyamine scaffold with uracil **5** and library aldehydes **6** through **34**. Mixtures that showed inhibitory activity were subjected to fractionation using reversed-phase HPLC, and then the individual purified components were reevaluated to determine which species was responsible for the inhibition. After identifying the inhibitory molecules, they were resynthesized in larger scale, purified, and complete IC₅₀ curves were determined as previously described.³³

Two RCHO groups were found to be inhibitory when attached to uracil in the triskelion scaffold (**Table 2**). The IC₅₀ values were found to fall in the range ~0.9 to 11 μM. When RCHO = 3,4 dihydroxybenzaldehyde, both the RRU (**45**) and UUR (**46**) variants showed nearly equal activity, suggesting that the enzyme recognizes the UR element, but not the third substituent on the scaffold (U or R). Consistent with this interpretation, the corresponding bifunctional oxime (UR, **47**, **Table 2**)³³ showed a similar IC₅₀ value as **45** and **46**. In contrast, when RCHO = 3-carboxybenzaldehyde, the trivalent forms **48** and **49** were found to be 6- to 12-fold more potent than the bifunctional oxime **50**. Thus, in this latter case the third substituent has a significant effect on binding affinity.

2.3. High-throughput screening of triskelion libraries against dUTPase

A similar strategy was used to screen the 215 oxime mixtures obtained from reaction of the triskelion scaffold with uracil aldehyde **4** and library aldehydes **6** through **34** against human dUTPase. In this case, only one inhibitory RCHO group was found (**Table 3**), and both the RRU (**51**) and UUR (**52**) variants provided comparable IC₅₀ values in the range 3–5 μM. To the best of our knowledge, **51** and **52** are the most potent nonnucleotide inhibitors of human dUTPase yet reported.

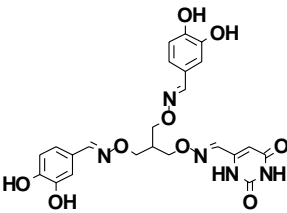
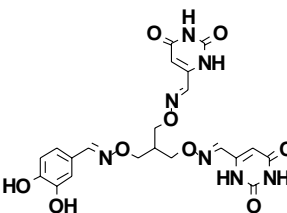
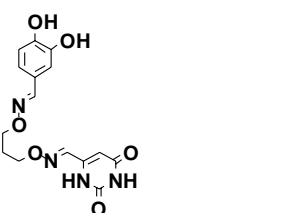
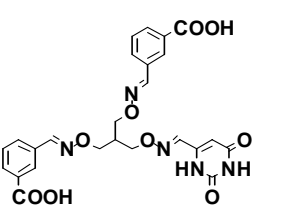
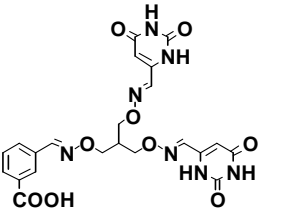
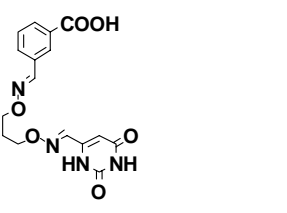
2.4. Inhibition by the untethered parts

The key question in determining the effectiveness of this tethering strategy is the inhibitory capacities of the untethered uracil and aldehyde components. For UNG2, the methyl oxime of uracil **5** has an IC₅₀ value of 75 μM, and no inhibition by the methyl oxime of 3,4 dihydroxybenzaldehyde (**15**) could be detected even at concentrations as high as 1 mM. Thus, tethering **15–5** produced an increase in binding affinity of 75-fold relative to uracil alone, and tethering **5–15** brought about at least a 1000-fold increase in binding affinity relative to **15** alone. For dUTPase, the methyl oximes of **4** and **8** were not inhibitory even at concentrations as high as 1 mM. Thus in this case, tethering of the two parts has brought about increases in binding affinity of at least 700-fold as compared to the separate components. A trivial but potentially useful modification of the tethering approach would be to incorporate two different R groups into the triskelion scaffold. This is easily accomplished by first synthesizing and isolating the monoderivatized uracil compound and then reacting the remaining two oxyamine positions with a mixture of two aldehydes (unpublished).

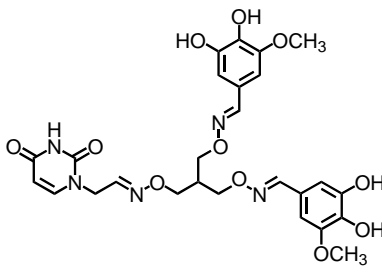
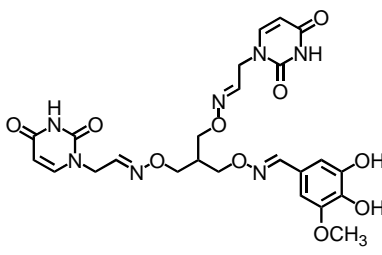
3. Conclusion

We have established that triskelion libraries of uracil derivatives can efficiently yield inhibitors with micromolar affinities for two different enzymes that recognize the

Table 2. Structures and inhibitory constants for UNG2 inhibitors^a

Triskelion compound	IC ₅₀ (μM)
 45	1.6 ± 0.1
 46	0.9 ± 0.1
 47	1.3 ± 0.1
 48	1.7 ± 0.5
 49	0.9 ± 0.1
 50	10.9 ± 0.5

^a See Ref. 33 for representative HTS screening data and IC₅₀ curves.**Table 3.** Structures and inhibitory constants for dUTPase inhibitors

Triskelion compound	IC ₅₀ (μM)
 51	3.3 ± 1.1
 52	5.5 ± 1.0

uracil base. Unlike previous nucleic acid-based inhibitors of UNG,^{34,35} and nucleotide-based inhibitors of dUTPase,³⁶ these library compounds are expected to be cell permeable. A useful extension of this approach is currently being developed where the length of the linker arms is varied. This modified approach is expected to generate more diverse libraries that allow more comprehensive probing of potential binding sites near the uracil pocket.

With respect to the in vivo utility of such oxime libraries, there are a number of currently used drugs with oxime functional groups: the selective serotonin reuptake inhibitor, fluvoxamine,^{37,38} the monobactam antibiotic, aztreonam,³⁸ and several preclinical antimicrobial drugs.^{39,40} The activity of these drugs indicates that oxime linkages are stable and useful in real clinical applications. Nevertheless, oximes are susceptible to reduction in metabolic reactions involving cytochrome P450-mediated transformations.^{40,41} Depending on the pharmacokinetic and pharmacodynamic properties of the individual oximes, this may or may not pose a problem. For instance fluvoxamine, although extensively processed in first-pass metabolism, has a reasonable serum half-life of 12 h.^{37,38} We anticipate that triskelion libraries based on substrate fragments will be useful for rapid inhibitor development against a variety of enzymes.

4. Experimental

4.1. Reagents and general methods

All chemicals were purchased from commercial sources without further purification unless otherwise stated. The ¹H and ¹³C NMR spectra were recorded on a 400 MHz Varian Innova instrument. The spectra were recorded in deuteriochloroform (CDCl₃) or in hexadeu-

teriodimethylsulfoxide (DMSO- d_6). The chemical shifts of protons are given in ppm with TMS as internal standard. The chemical shifts of carbons are obtained in ppm with solvents as internal standards. Oximes were purified by HPLC using aqueous triethylammonium acetate (TEAA) as a running buffer. Therefore, TEAA was not completely removed and it appeared in the NMR spectra. Accordingly, proton and carbon chemical shifts of TEAA were not listed during the characterizations of the oximes. During purification of the oximes **45** and **46**, 2-mercaptoethanol was used as an anti-oxidant. Therefore, small amounts of this compound and its oxidation product are also present in these oximes. Flash chromatographies were performed with silica (70–230 mesh from Sorbent Technologies) and monitored by thin-layer chromatography (TLC) with silica plates (Merck, Kieselgel 60 F254).

4.2. 2,2',2''-[(2-Methyloxy-1,3-propanedioxy)tris]-1*H*-isoindole-1, 3(2*H*)-dione (**2**)

To a suspension of 2-hydroxymethyl-1,3-propanediol **1** (0.848 g, 8.0 mmol), triphenylphosphine (7.32 g, 28 mmol), and *N*-hydroxyphthalimide (6.52 g, 40 mmol) in anhydrous THF (60 ml) was added diisopropyl azodicarboxylate (5.64 ml, 28.0 mmol) dropwise at 0 °C.^{42–44} The mixture was stirred overnight, and the precipitate was filtered and washed with cold THF. After removal of the THF in vacuo, product **2** was obtained (2.3 g) in 53% yield. ¹H NMR (400 MHz, CDCl₃): δ 7.76 (m, 12H), 4.66 (d, J = 6.4 Hz, 6H), 2.78 (septet, J = 6.4 Hz, 1H); ¹³C NMR (100 MHz, CDCl₃): δ 163.43, 134.39, 128.97, 123.47, 75.52, 37.68. HRMS (m/z): [M+H]⁺ calcd for C₂₈H₂₀N₃O₉, 542.12; found, 542.12.

4.3. 2-[(Aminooxy)-methyl]-1,3-bis(aminooxy)-propane (**3**)

To a suspension of **2** (2.43 g, 4.5 mmol) in 95% ethanol (9.5 ml), was added anhydrous hydrazine (0.67 ml, 20.2 mmol) dropwise within 10 min at room temperature. The mixture was stirred for 2 h, filtered, and washed with 95% ethanol. The filtrate was concentrated in vacuo to give a residue. To the residue was added methylene chloride (10 ml). The resulting mixture was kept overnight at room temperature, filtered the following morning, and the filtrate was concentrated in vacuo, giving product **3** (0.46 g) in 67% yield. ¹H NMR (400 MHz, CDCl₃): δ 3.71 (d, J = 6.0 Hz, 6H), 2.41 (septet, J = 6.4 Hz, 1H); ¹³C NMR (100 MHz, CDCl₃): δ 74.66, 37.05; HRMS (m/z): [M+H]⁺ calcd for C₄H₁₄N₃O₃, 152.10; found, 152.10.

4.4. General synthesis of tethered triskelion oximes

A set of 39 commercially available aldehydes were linked to the two uracil-containing aldehydes **4** and **5** using the triskelion alkyloxyamine linker **3** (see Scheme 2 and Table 1, and supplementary materials). To each well of a 0.5-ml Matrix microtiter plate were added DMSO stock solutions of AcOH (20 μ L, 150 mM, 3 μ mol), uracil-containing aldehyde **4** or **5** (20 μ L, 150 mM, 3 μ mol), and one library aldehyde (40 μ L, 150 mM, 6 μ mol). The synthesis and characterization

of **4** and **5** has been previously reported. The plate was carefully agitated to make the solutions homogeneous, and 22 μ L of a DMSO solution of the triskelion oxyamine was then added (150 mM, 3.3 μ mol). The plate was sealed and further agitated and incubated in an oven for 12 h at 37 °C. The expected statistical ratios of the oxime products were confirmed by ¹H NMR analysis (see supplementary materials).

4.5. Isolation and purification of inhibitory triskelion oximes

The most potent inhibitors were resynthesized in larger scale and thoroughly characterized after HPLC purification of the mixed oximes using a Phenomenex Aqua reversed-phase C-18 HPLC column (250 mm, 10 mm, 5 μ m). Gradient elution from 0% to 65% CH₃CN in 0.1 M aqueous TEAA over the course of 2 h with UV detection at 254 nm was used. An exception was oxime **47**, which is prone to air oxidation. In this case, 25 mM 2-mercaptoethanol was added to both of the running buffers. The oximes all eluted with baseline resolution in the order (1) the homotrimer oxime derived from **4** or **5**, (2) the heterotrimer oxime derived from **4** or **5** and either **8**, **15** or **19**, and (3) the homotrimer oxime derived from **8**, **15** or **19**.

4.6. 2-[*O*-(6-Uracilcarboxaldoximyl)-methyl]-1,3-bis[*O*-(3,4-dihydroxybenzaldoximyl)]-propane (**45**)

¹H NMR (400 MHz, DMSO- d_6): δ 11.15 (br s, 1H), 9.30 (br s, 1H), 8.09 (s, 2H), 7.98 (s, 1H), 7.05 (d, J = 1.6 Hz, 1H), 6.84 (dd, J = 8.0, 2.0 Hz, 1H), 6.74 (d, J = 8.0 Hz, 1H), 5.78 (s, 1H), 4.32 (d, J = 6.0 Hz, 2H), 4.17 (d, J = 5.6 Hz, 4H), 2.66 (m, 1H); ¹³C NMR (100 MHz, DMSO- d_6): δ 163.85, 151.04, 149.12, 147.70, 145.57, 144.51, 142.43, 123.09, 119.95, 115.60, 112.91, 101.78, 73.25, 71.13, 38.67; HRMS (m/z): [M+H]⁺ calcd for C₂₃H₂₄N₅O₉, 514.16; found, 514.16.

4.7. 2-[*O*-(3,4-Dihydroxybenzaldoximyl)-methyl]-1,3-bis[*O*-(6-uracilcarboxaldoximyl)]-propane (**46**)

¹H NMR (400 MHz, DMSO- d_6): δ 11.15 (br s, 1H), 8.05 (s, 1H), 7.97 (s, 2H), 7.04 (d, J = 2.0 Hz, 1H), 6.84 (dd, J = 8.0, 2.0 Hz, 1H), 6.74 (d, J = 8.0 Hz, 1H), 5.78 (s, 2H), 4.32 (d, J = 6.4 Hz, 4H), 4.16 (d, J = 6.0 Hz, 2H), 2.66 (m, 1H); ¹³C NMR (100 MHz, DMSO- d_6): δ 164.50, 151.72, 149.84, 148.39, 146.24, 145.16, 143.13, 123.68, 120.61, 116.25, 113.61, 102.40, 73.73, 71.60, 38.98; HRMS (m/z): [M+H]⁺ calcd for C₂₁H₂₂N₇O₉, 516.15; found, 516.15.

4.8. 1-[*O*-(6-Uracilcarboxaldoximyl)]-3-[*O*-(3,4-dihydroxybenzaldoximyl)]-propane (**47**)

The synthesis and characterization of this oxime heterodimer has been previously described³³.

4.9. 2-[*O*-(6-Uracilcarboxaldoximyl)-methyl]-1,3-bis[*O*-(3-carboxybenzaldoximyl)]-propane (**48**)

¹H NMR (400 MHz, DMSO- d_6): δ 8.34 (s, 2H), 8.15 (s, 2H), 8.0 (s, 1H), 7.91 (d, J = 6.8 Hz, 2H), 7.65 (d,

$J = 8.0$ Hz, 2H), 7.38 (t, $J = 8.4$ Hz, 2H), 5.77 (s, 1H), 4.36 (d, $J = 6.4$ Hz, 2H), 4.28 (d, $J = 5.6$ Hz, 4H), 2.82 (m, 1H); ^{13}C NMR (100 MHz, DMSO- d_6): δ 168.54, 163.88, 151.07, 149.16, 144.57, 142.60, 137.36, 131.43, 130.62, 128.75, 128.27, 127.55, 101.86, 73.12, 71.50, 38.67; HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{25}\text{H}_{24}\text{N}_5\text{O}_9$, 538.16; found, 538.16.

4.10. 2-[*O*-(3-Carboxybenzaldoximyl)-methyl]-1,3-bis[*O*-(6-uracilcarboxaldoximyl)]-propane (49)

^1H NMR (400 MHz, DMSO- d_6): δ 11.17 (br s, 1H), 8.32 (s, 1H), 8.16 (s, 1H), 8.0 (s, 2H), 7.92 (dd, $J = 6.8$, 1.6 Hz, 1H), 7.65 (d, $J = 8.0$ Hz, 1H), 7.39 (t, $J = 7.2$ Hz, 1H), 5.79 (s, 2H), 4.33 (d, $J = 5.6$ Hz, 4H), 4.25 (d, $J = 5.6$ Hz, 2H), 2.82 (m, 1H); ^{13}C NMR (100 MHz, DMSO- d_6): δ 168.29, 163.85, 151.10, 149.15, 144.54, 142.46, 131.33, 130.67, 128.22, 127.58, 101.70, 72.99, 71.39, 38.56; HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{22}\text{H}_{22}\text{N}_7\text{O}_9$, 528.15; found, 528.15.

4.11. 1-[*O*-(6-Uracilcarboxaldoximyl)]-3-[*O*-(3-carboxybenzaldoximyl)]-propane (50)

The synthesis and characterization of this oxime heterodimer has been previously described³³.

4.12. 2-{*O*-[2-(N1-uracil)-acetaldoximyl]-methyl}-1,3-bis[*O*-(3,4-dihydroxy-5-methoxybenzaldoximyl)]-propane (51)

^1H NMR (400 MHz, DMSO- d_6) δ 11.32 (d, 2H), 9.14 (s, 1H), 8.72 (s, 1H), 8.03 (q, 1H), 7.65 (dd, 1H), 7.56 (d, 1H), 7.52 (dd, 1H), 6.92 (dd, 1H), 6.69 (s, 1H), 5.55 (d, 2H), 4.50 (dd, 2H), 4.43 (t, 2H), 4.17–4.00 (m, 6H), 3.70 (s, 3H), 1.28 (m, 1H); ^{13}C NMR (400 MHz, DMSO- d_6) δ (164.5, 164.4), 153.2, (151.8, 151.5), (150.8, 150.0), (149.1, 147.2), (146.5, 146.3), 143.5, 137.1, 122.7, 108.6, 103.2, (101.9, 101.8), 72.4, 72.0, 59.9, (56.9, 56.5), (46.5, 45.2); HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{24}\text{H}_{28}\text{N}_7\text{O}_{10}$ 574.1892, found 574.1885.

4.13. 2-[*O*-(3,4-Dihydroxy-5-methoxybenzaldoximyl)-methyl]-1,3-bis[*O*-(2-(N1-uracil)-acetaldoximyl)]-propane (52)

^1H NMR (400 MHz, DMSO- d_6) δ 11.33 (d, 1H), 9.14 (s, 2H), 8.72 (s, 2H), 8.03 (d, 2H), 7.65 (d, 0.5H), 7.56 (d, 0.5H), 7.53 (t, 0.5H), 6.93 (t, 0.5H), 6.69 (s, 2H), 5.55 (d, 1H), 4.50 (d, 1H), 4.43 (d, 1H), 4.42–4.00 (m, 6H), 3.70 (s, 6H), 1.24 (m, 1H); ^{13}C NMR (400 MHz, DMSO- d_6) δ (164.5, 164.4), 153.2, (151.8, 151.5), (150.8, 150.0), (149.1, 147.2), (146.5, 146.3), 143.5, 137.1, 122.7, 108.7, 103.1, (101.9, 101.8), 72.4, 71.9, 59.9, (56.9, 56.4), (46.5, 45.2); HRMS (m/z): $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{26}\text{H}_{30}\text{N}_5\text{O}_{11}$ 588.1936, found 588.1920.

5. In vitro inhibition studies

5.1. High-throughput inhibitor screening of UDG

The molecular beacon-based HTS assay for UDG has been previously described.³³ Briefly, to a 96-well micro-

titer plate was added 5 μL of 2 mM total compound in DMSO, followed by 75 μL 33.3 pM human UNG in reaction buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 7.5 mM MgCl_2 , 0.002% brij-35). The reactions were initiated by the addition of 20 μL of 250 nM molecular beacon substrate in reaction buffer. The plates are incubated at ambient temperature in a fluorescence plate reader for 30 min, and the progress of the reaction was monitored every 5 min (Ex. 485 nm/Em. 520 nm). The final concentrations of the reagents in the assay are 10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 7.5 mM MgCl_2 , 0.002% Brij-35, 25 or 100 pM human UNG, 50 nM molecular beacon substrate, 100 μM total compound, and 5% DMSO. IC_{50} analysis was performed using the same conditions except that the concentration of compound was varied in the range 0.01–100 μM .

5.2. High-throughput inhibitor screening of dUTPase

To a 96-well microtiter plate were added 20 μL of reaction buffer (50 mM Tris, pH 8.0, 10 mM MgCl_2 , and 0.05% Tween 20) and 5 μL of compounds (2 mM) in DMSO, followed by 50 μL of dUTPase (1 nM) and pyrophosphatase (10 U/mL) in reaction buffer. The reactions were initiated by the addition of 25 μL dUTP (40 μM). The plates are incubated at ambient temperature for 50 min and the reactions were quenched by 25 μL of malachite green color reagent.⁴⁵ The mixtures were then allowed to stand for 10 min and the absorbances were measured using a microtiter plate reader with a 620 nm bandpass filter. IC_{50} analysis was performed using the same conditions except that the concentration of compound was varied in the range 0.1–75 μM .

Acknowledgments

This work was supported by NIH Grant GM56834-10 to J.T.S. D.J.K. was supported by the DOD Breast Cancer Research Program (DAMD17-03-1-1251).

Supplementary data

Tables of aryl aldehydes and hydroxybenzaldehydes used to construct oxime libraries, NMR confirmation of product ratios for representative oximes, and IC_{50} curves for active compounds. Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bmc.2006.04.022](https://doi.org/10.1016/j.bmc.2006.04.022).

References and notes

- Lindahl, T.; Wood, R. D. *Science* **1999**, *286*, 1897–1905.
- Unniraman, S.; Fugmann, S. D.; Schatz, D. G. *Science* **2004**, *305*, 1113–1114.
- Seiple, L.; Jaruga, P.; Dizdaroglu, M.; Stivers, J. T. *Nucleic Acids Res.* **2006**, *34*, 140–151.
- Saribasak, H.; Saribasak, N. N.; Ipek, F. M.; Ellwart, J. W.; Arakawa, H.; Buerstedde, J. M. *J. Immunol.* **2006**, *176*, 365–371.
- Tye, B. K.; Lehman, I. R. *J. Mol. Biol.* **1977**, *117*, 293–306.

6. Warner, H. R.; Duncan, B. K. *Nature* **1978**, 272, 32–34.
7. Nilsen, H.; Haushalter, K. A.; Robins, P.; Barnes, D. E.; Verdine, G. L.; Lindahl, T. *EMBO J.* **2001**, 20, 4278–4286.
8. Di Noia, J.; Neuberger, M. S. *Nature* **2002**, 419, 43–48.
9. Imai, K.; Slupphaug, G.; Lee, W. I.; Revy, P.; Nonoyama, S.; Catalan, N.; Yel, L.; Forveille, M.; Kavli, B.; Krokan, H. E.; Ochs, H. D.; Fischer, A.; Durandy, A. *Nat. Immunol.* **2003**, 4, 1023–1028.
10. Storb, U.; Stavnezer, J. *Curr. Biol.* **2002**, 12, R725–R727.
11. Chen, R.; Wang, H.; Mansky, L. M. *J. Gen. Virol.* **2002**, 83, 2339–2345.
12. Prichard, M. N.; Duke, G. M.; Mocarski, E. S. *J. Virol.* **1996**, 70, 3018–3025.
13. De Silva, F. S.; Moss, B. *J. Virol.* **2003**, 77, 159–166.
14. Stuart, D. T.; Upton, C.; Higman, M. A.; Niles, E. G.; McFadden, G. *J. Virol.* **1993**, 67, 2503–2512.
15. Priet, S.; Gros, N.; Navarro, J. M.; Boretto, J.; Canard, B.; Querat, G.; Sire, J. *Mol. Cell* **2005**, 17, 479–490.
16. Ladner, R. D. *Curr. Protein Pept. Sci.* **2001**, 2, 361–370.
17. Tinkelenberg, B. A.; Hansbury, M. J.; Ladner, R. D. *Cancer Res.* **2002**, 62, 4909–4915.
18. Stivers, J. T.; Drohat, A. C. *Arch. Biochem. Biophys.* **2001**, 396, 1–9.
19. Bouhamdan, M.; Benichou, S.; Rey, F.; Navarro, J. M.; Agostini, I.; Spire, B.; Camonis, J.; Slupphaug, G.; Vigne, R.; Benarous, R.; Sire, J. *J. Virol.* **1996**, 70, 697–704.
20. BouHamdan, M.; Xue, Y.; Baudat, Y.; Hu, B.; Sire, J.; Pomerantz, R. J.; Duan, L. X. *J. Biol. Chem.* **1998**, 273, 8009–8016.
21. Chen, R.; Le Rouzic, E.; Kearney, J. A.; Mansky, L. M.; Benichou, S. *J. Biol. Chem.* **2004**, 279, 28419–28425.
22. Klarmann, G. J.; Chen, X.; North, T. W.; Preston, B. D. *J. Biol. Chem.* **2003**, 278, 7902–7909.
23. Mansky, L. M.; Preveral, S.; Selig, L.; Benarous, R.; Benichou, S. *J. Virol.* **2000**, 74, 7039–7047.
24. Payne, S. L.; Elder, J. H. *Curr. Protein Pept. Sci.* **2001**, 2, 381–388.
25. Selig, L.; Benichou, S.; Rogel, M. E.; Wu, L. I.; Vodicka, M. A.; Sire, J.; Benarous, R.; Emerman, M. *J. Virol.* **1997**, 71, 4842–4846.
26. Willetts, K. E.; Rey, F.; Agostini, I.; Navarro, J. M.; Baudat, Y.; Vigne, R.; Sire, J. *J. Virol.* **1999**, 73, 1682–1688.
27. Elder, R. T.; Zhu, X.; Priet, S.; Chen, M.; Yu, M.; Navarro, J. M.; Sire, J.; Zhao, Y. *Biochem. Biophys. Res. Commun.* **2003**, 306, 693–700.
28. Mansky, L. M.; Le Rouzic, E.; Benichou, S.; Gajary, L. C. *J. Virol.* **2003**, 77, 2071–2080.
29. Miller, R. J.; Cairns, J. S.; Bridges, S.; Sarver, N. *J. Virol.* **2000**, 74, 7187–7195.
30. Lau, A.; Swinbank, K. M.; Ahmed, P. S.; Taylor, D. L.; Jackson, S. P.; Smith, G. C.; O'Connor, M. J. *Nat. Cell Biol.* **2005**, 7, 493–500.
31. Nilsen, H.; Rosewell, I.; Robins, P.; Skjelbred, C. F.; Andersen, S.; Slupphaug, G.; Daly, G.; Krokan, H. E.; Lindahl, T.; Barnes, D. E. *Mol. Cell* **2000**, 5, 1059–1065.
32. Maier, L.; Spoerri, H. Organic phosphorus compounds. 103. Aminoalkylphosphonic acids and aminoalkylphosphine oxides and derivatives. In *Phosphorus, Sulfur and Silicon and the Related Elements*, Ed.; Vol. 70; 1992; pp 49–57.
33. Jiang, Y. L.; Krosky, D. J.; Seiple, L.; Stivers, J. T. *J. Am. Chem. Soc.* **2005**, 127, 17412–17420.
34. Jiang, Y. L.; Cao, C.; Stivers, J. T.; Song, F.; Ichikawa, Y. *Bioorg. Chem.* **2004**, 32, 244–262.
35. Krosky, D. J.; Song, F.; Stivers, J. T. *Biochemistry* **2005**, 44, 5949–5959.
36. Whittingham, J. L.; Leal, I.; Nguyen, C.; Kasinathan, G.; Bell, E.; Jones, A. F.; Berry, C.; Benito, A.; Turkenburg, J. P.; Dodson, E. J.; Ruiz Perez, L. M.; Wilkinson, A. J.; Johansson, N. G.; Brun, R.; Gilbert, I. H.; Gonzalez Pacanowska, D.; Wilson, K. S. *Structure* **2005**, 13, 329–338.
37. Spigset, O.; Granberg, K.; Hagg, S.; Soderstrom, E.; Dahlqvist, R. *Br. J. Clin. Pharmacol.* **1998**, 45, 257–263.
38. Hemeryck, A.; Belpaire, F. M. *Curr. Drug Metab.* **2002**, 3, 13–37.
39. Sturk, L. M.; Brock, J. L.; Bagnell, C. R.; Hall, J. E.; Tidwell, R. R. *Acta Trop.* **2004**, 91, 131–143.
40. Heberling, S.; Girreser, U.; Wolf, S.; Clement, B. *Biochem. Pharmacol.* **2006**, 71, 354–365.
41. Kurian, J. R.; Bajad, S. U.; Miller, J. L.; Chin, N. A.; Trepanier, L. A. *J. Pharmacol. Exp. Ther.* **2004**, 311, 1171–1178.
42. Maly, D. J.; Choong, I. C.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97, 2419–2424.
43. Kung, P. P.; Bharadwaj, R.; Fraser, A. S.; Cook, D. R.; Kawasaki, A. M.; Cook, P. D. *J. Org. Chem.* **1998**, 63, 1846–1852.
44. Weiss, R. H.; Furfine, E.; Hausleden, E.; Dixon, D. W. *J. Org. Chem.* **1984**, 49, 4969–4972.
45. Geladopoulos, T. P.; Sotiroidis, T. G.; Evangelopoulos, A. E. *Anal. Biochem.* **1991**, 192, 112–116.

Mimicking damaged DNA with a small molecule inhibitor of human UNG2

Daniel J. Krosky, Mario A. Bianchet¹, Lauren Seiple, Suhman Chung,
L. Mario Amzel¹ and James T. Stivers*

Department of Pharmacology and Molecular Sciences and ¹Department of Biophysics and Biophysical Chemistry of the Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, MD 21205, USA

Received September 5, 2006; Revised September 22, 2006; Accepted September 26, 2006

ABSTRACT

Human nuclear uracil DNA glycosylase (UNG2) is a cellular DNA repair enzyme that is essential for a number of diverse biological phenomena ranging from antibody diversification to B-cell lymphomas and type-1 human immunodeficiency virus infectivity. During each of these processes, UNG2 recognizes uracilated DNA and excises the uracil base by flipping it into the enzyme active site. We have taken advantage of the extrahelical uracil recognition mechanism to build large small-molecule libraries in which uracil is tethered via flexible alkane linkers to a collection of secondary binding elements. This high-throughput synthesis and screening approach produced two novel uracil-tethered inhibitors of UNG2, the best of which was crystallized with the enzyme. Remarkably, this inhibitor mimics the crucial hydrogen bonding and electrostatic interactions previously observed in UNG2 complexes with damaged uracilated DNA. Thus, the environment of the binding site selects for library ligands that share these DNA features. This is a general approach to rapid discovery of inhibitors of enzymes that recognize extrahelical damaged bases.

INTRODUCTION

The RNA base uracil is one of the most prevalent non-canonical bases found in genomic DNA (1). It arises from spontaneous or intentional enzymatic deamination of cytosine in DNA (2–5), or alternatively, by misincorporation of dUTP in place of TTP during DNA replication (6). Both pathways for uracil incorporation are forms of DNA damage, and accordingly, an elaborate uracil base excision repair (UBER) mechanism is present in all organisms to reverse this damage (Figure 1A) (7). Without repair, U/G mismatches lead to T/A transition mutations and corresponding changes in protein sequence. Although U/A base pairs arising from

misincorporation of dUTP are not mutagenic, if large numbers of uracils are inserted on both strands of replicated DNA this can lead to disruptions in gene expression, and even double strand DNA breaks can arise from the base excision repair process (8). Although the accidental appearance of uracil in DNA is well-appreciated, it has become apparent that enzymatic deamination of cytosine to uracil in DNA plays a key role in the processes of somatic hypermutation and class switch recombination in B cells (2,9,10), in certain B cell lymphomas (11), and as an innate host defense mechanism against retroviral infection (12). In addition, the widely used chemotherapeutic agent 5-fluorouracil (5-FU) promotes uracil misincorporation into DNA by increasing the ratio [dUTP]/[TTP] in the cell, suggesting inhibitors of UBER could serve as sensitizers during 5-FU therapy (8,13). In general, these diverse roles for uracil indicate that small molecule UBER inhibitors might be very useful investigational or therapeutic agents.

In humans, the first step in the UBER pathway, cleavage of the glycosidic bond of deoxyuridine in DNA, is catalyzed by the powerful nuclear enzyme uracil DNA glycosylase (UNG2) (14). This extensively studied glycosylase uses an extrahelical recognition mechanism in which the uracil base that is originally embedded in the DNA base stack is ultimately extruded into the enzyme active site (Figure 1B) (15). Thus by ‘uracil flipping’ the enzyme can recognize the unique structural features of uracil that allows catalysis to proceed (7). Although the uracil is attached to a large duplex DNA substrate, most of the enzyme–substrate interactions involve the base itself. Perhaps not surprisingly, the uracil base alone has been found to be a product inhibitor of the enzyme ($K_i \sim 300 \mu\text{M}$ at physiological pH) (16).

We recently published a general strategy for rapid discovery of small molecule inhibitors of UNG2 and other UBER enzymes called ‘substrate fragment tethering’ (SFT) (17), which is an efficient variation of the combinatorial target-guided ligand assembly method of Ellman *et al.* (18). The basic approach is extremely simple and involves tethering a chemical library of aldehydes to pieces of substrates (such as uracil) that already bind weakly to an enzyme active site. As shown in Figure 2, UNG2 library synthesis involves

*To whom correspondence should be addressed. Tel: +1 410 502 2758; Fax: +1 410 955 3023; Email: jstivers@jhmi.edu

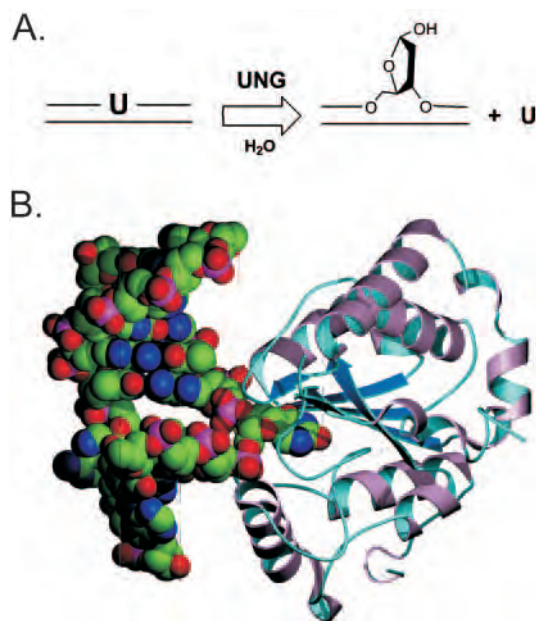


Figure 1. Uracil DNA base excision repair and extrahelical recognition of uracil. (A) Uracil in the context of a U/A or U/G base pair is repaired by a series of enzymatic reactions that restore the integrity of the DNA sequence. The first enzyme in the pathway is uracil DNA glycosylase (UNG) that hydrolytically cleaves the *N*-glycosidic bond connecting the uracil to the deoxyribose, leaving an abasic site and free uracil. Three other enzymes complete the repair process in humans: an abasic site endonuclease (APE1), a dual-function repair polymerase (pol β) that inserts the correct nucleotide and eliminates the abasic site via a β -elimination reaction, and finally, DNA ligase. (B) Structure of uracilated DNA bound to human UNG2 (PDB code 1EMH).

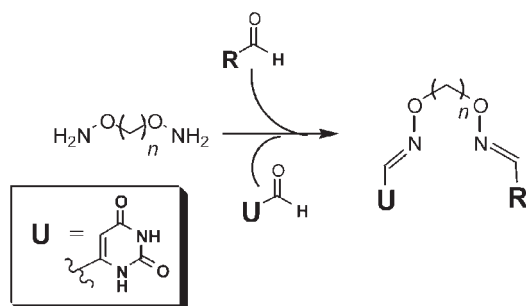


Figure 2. Chemistry of substrate fragment tethering. In this approach, a suitable substrate fragment (such as uracil) is identified and derivatized at a nonperturbing position with an aldehyde functional group. In the case of UNG2, the substrate fragment is 6-formyluracil. The substrate fragment is tethered to one end of a bifunctional alkyloxyamine linker of variable length ($n = 2-6$), which is then derivatized on the other end with a library of aldehyde binding elements (RCHO). Although statistical mixtures of all possible oximes result (25% each homodimer derived from uracil-CHO or RCHO, and 50% of the heterodimer derived from uracil-CHO and RCHO), this poses no difficulty because the crude mixtures are directly screened for inhibitory activity. Once active mixtures are identified, the specific inhibitory components can be rapidly identified by deconvolution to uncover the linker length that gave rise to the observed inhibition. The structures of the 215 aldehyde library members used in this study are reported in Supplementary Table S1.

highly efficient formation of oxime linkages between a bivalent alkyloxyamine linker, a uracil aldehyde derivative, and each library aldehyde member. Thus, the uracil fragment targets the entire tethered molecule to the active site where the

library pieces can then explore adjacent binding pockets. SFT has the following strengths: (i) library synthesis is economical and very rapid and can be performed in microtiter plate format, (ii) the reactions are extremely efficient and no purification of any products is required, (iii) mixtures of flexible linkers are used in each reaction which allows multiple tethering lengths to be probed simultaneously in activity screens, and (iv) the method is easily adaptable to any desired target.

Here we report the results from high-throughput screening (HTS) of an SFT library derived from tethering 6-formyluracil to a library of 215 different aldehyde-binding elements. Of the two hits identified in this screen, the most potent SFT ligand was co-crystallized with UNG2 to yield a high-resolution structure of the complex. This first portrait of a bound SFT ligand shows how a small molecule can surprisingly mimic the hydrogen bonding and electrostatic interactions of a larger DNA substrate. Thus, the environment of the binding site appears to select for library ligands that share molecular features of DNA. This efficient approach should be easily adaptable to other DNA repair glycosylases that recognize extrahelical bases.

MATERIALS AND METHODS

Reagents and general methods

All chemicals were purchased from commercial sources without further purification unless otherwise stated. The ^1H NMR spectra were recorded on a 400 MHz Varian Innova instrument in hexadeuteriodimethylsulfoxide (DMSO-d_6). The chemical shifts of protons are given in p.p.m. with DMSO as an internal standard.

Synthesis of tethered oxime libraries

The 6-formyluracil tethered library was synthesized as described previously (17). Briefly, a set of 215 alkyl and aryl aldehydes (Supplementary Table S1) was selected for library synthesis and coupled to 6-formyluracil using the *O,O'*-diaminoalkanedil linkers as follows. To each 0.5 ml well of a Matrix microtiter plate was added a DMSO stock solution of AcOH (20 μl , 150 mM, 3 μmol), 6-formyluracil (20 μl , 150 mM, 3 μmol) and a single alkyl or aryl aldehyde (20 μl , 150 mM, 3 μmol). The plate was carefully agitated to make the solutions homogenous. To each of the mixtures was added a DMSO solution of the *O,O'*-diaminoalkanedil linkers containing each of the five linker lengths in equal proportion (22 μl , 150 mM, 3.3 μmol total amine equivalents). The plate was sealed, further agitated and incubated in an oven for 12 h at 37°C.

Deconvolution of inhibitory mixtures

The two active mixtures containing compounds **1** and **2** were deconvoluted with respect to linker length by individually synthesizing each oxime dimer using a *single* diaminoalkanedil linker per reaction. At this stage we did not separate the homodimers from the heterodimers in the mixtures. The linker length dependence of the inhibition is reported in Supplementary Table S2. The corresponding compounds were then synthesized in larger scale and purified for complete analysis of their inhibition properties as described below.

Synthesis of 1

Solutions (0.15 M) of 4-carboxybenzaldehyde (888 μ l, 0.165 mmol), 6-formyluracil (888 μ l, 0.165 mmol) and acetic acid (888 μ l, 0.165 mmol) in DMSO were added to a reaction vessel. The reaction was initiated by the addition of 0.15 M *O,O'*-diaminoethanediol (888 μ l, 0.165 mmol) in DMSO, and incubated at 37°C for 36 h. The desired heterosubstituted compound was purified by direct injection of the reaction mixture onto a Phenomenex Aqua reversed phase C-18 HPLC column (250 mm, 10 mm, 5 μ m) using gradient elution from 0 to 65% CH₃CN in 0.1 M aqueous TEAA over the course of 2 h using UV detection at 320 nm. Fractions containing **1** were combined and concentrated *in vacuo*. The compound was precipitated using ice-cold water, centrifuged, washed twice with ice-cold water and dried *in vacuo*. This yielded **1** as a white powder (9.6 mg, 0.028 mmol) in 34% yield. ¹H NMR (400 MHz, DMSO-*d*₆): δ 11.19 (s, 1H), 10.81 (s, 1H), 8.37 (s, 1H), 7.98 (s, 1H), 7.96 (d, 3H), 7.72 (d, 2H), 5.79 (s, 1H), 4.66 (d, 4H); ¹³C NMR (400 MHz, DMSO-*d*₆): δ 167.6, 163.9, 151.0, 148.7, 144.5, 142.5, 129.8, 126.9, 102.0, 73.2, 71.9; (*m/z*): [M+H]⁺ calcd for C₁₅H₁₄N₄O₆, 347.0986; found, 347.0991.

Synthesis of 2

The synthesis and purification was identical to that reported above for **1** except that 0.15 M solutions of 3-carboxybenzaldehyde and *O,O'*-diaminopropanediol were used. ¹H NMR (400 MHz, DMSO-*d*₆): δ 11.16 (s, 1H), 10.78 (s, 1H), 8.34 (s, 1H), 8.18 (s, 1H), 7.94 (d, 2H), 7.83 (d, 2H), 7.53 (t, 1H), 5.76 (s, 1H), 4.27 (t, 2H), 4.20 (t, 2H), 2.10 (m, 2H); ¹³C NMR (400 MHz, DMSO-*d*₆): δ 166.9, 163.9, 151.0, 148.3, 144.6, 142.2, 132.5, 131.5, 130.9, 130.5, 129.2, 127.6, 101.7, 71.8, 70.5, 28.4; (*m/z*): [M+H]⁺ calcd for C₁₅H₁₄N₄O₆, 361.1143; found, 361.1153.

High-throughput screening of oxime library

The DNA substrate in this HTS assay was synthesized using standard phosphoramidite DNA solid-phase chemistry using reagents purchased from Glen Research. The DNA was purified using anion exchange chromatography followed by desalting using reversed phase methods. The sequence and size was confirmed using analytical denaturing PAGE and MALDI-MS. The substrate is a single-stranded 28mer DNA hairpin containing nine U-A base pairs and a hexapolyethylene glycol (PEG₆) linker (5'-FAM-GCA CUU AAG AAU UG-PEG₆-CA AUU CUU AAG UGC-DABSYL-3'). The UNG2 HTS assay has been described previously (17).

IC₅₀ determinations

To a 96-well plate was added 5 μ l compound **1** in DMSO, followed by 75 μ l of 66.5 nM PEG-U9 hairpin in reaction buffer (20 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂ and 0.05% Brij-35). Eight different inhibitor concentrations were used in the range of 0.045–100 μ M. Reactions were initiated by the addition of 20 μ l of 0.5 nM human UNG in reaction buffer. The final concentrations of reagents in the assay are 20 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂, 0.05% Brij-35, 5% DMSO, 0.1 nM human UNG, 50 nM PEG-U9 hairpin DNA and 0–100 μ M **1**. Wells containing DMSO vehicle only or no UNG2 were used as negative controls and background, respectively. The plates are incubated

at ambient temperature in a fluorescence plate reader for 30 min, and the progress of the reaction was monitored every 5 min (Ex. 485 nm/Em. 520 nm). Percent inhibition versus log concentration of **1** data were fit to a four parameter sigmoidal dose–response equation (Equation 1) using Prism 4.0 (GraphPad Software Inc., San Diego, CA, USA).

$$\% \text{ Inhibition} = \frac{\text{Min}(\text{Max} - \text{Min})}{1 + 10^{((\log \text{IC}_{50} - \log[1]) * n)}} \quad 1$$

Mechanism of inhibition

To a 96-well plate was added 5 μ l compound in DMSO, followed by 75 μ l PEG-U9 hairpin in reaction buffer (20 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂ and 0.05% Brij-35). Eight different DNA concentrations were used in the range of 27.5–1100 nM. Reactions were initiated by the addition of 20 μ l of 0.5 nM human UNG in reaction buffer. The final concentrations of reagents in the assay are 20 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.2 mM MgCl₂, 0.05% Brij-35, 5% DMSO, 0.1 nM human UNG, 27.5–1100 nM PEG-U9 hairpin DNA and 0–128 μ M of **1**. The plates were incubated at ambient temperature in a fluorescence plate reader for 60 min, and the progress of each reaction was monitored every 30 s (λ_{ex} = 485 nm, λ_{em} = 520 nm). Afterwards, *Escherichia coli* UNG was added to each well to drive the reactions to completion, and the total change in fluorescence corresponding to complete consumption of the substrate was calculated ($\Delta \text{FU}_{\text{tot}}$). These values were used to calculate initial molar velocities (i.e. [product]/s = $\Delta \mu\text{M} / \Delta \text{FU}_{\text{tot}} \times \text{FU/s}$). Mechanisms of inhibition and their corresponding inhibitor dissociation constants were determined by Lineweaver–Burk slope and intercept replot analysis.

Cell culture studies

Currently, there exists no simple assay to monitor the efficacy of inhibitors of UNG within living cells. In order to evaluate the ability of compound **1** to inhibit UNG *in vivo*, inhibitor was added to cells in culture, lysates were carefully prepared to minimize dilution of the inhibitor, and UNG activity was assayed as follows (Figure 3). PC-3 human prostate adenocarcinoma cells were grown in Modified Eagle's Medium (Mediatech) supplemented with 10% fetal bovine serum (Hyclone) and penicillin/streptomycin. Cells at 70% confluency were trypsinized, resuspended in growth media and counted using a hemocytometer. The cells were then centrifuged and resuspended to a density of 1×10^6 cells/ml. Aliquots of 1×10^6 cells were pelleted and resuspended in 46.5 μ l growth media with or without 3 mM compound **1** to achieve a final concentration of 3 mM inhibitor. The cells were incubated at 37 °C, 5% CO₂ for 30 min before being spun down and resuspended in 7 μ l UNG lysate reaction buffer [10 mM Tris-HCl, pH 8.0, 60 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.002% Brij 35, protease inhibitors (Roche Complete, Mini)]. The cells were lysed by freeze-thawing five times, the lysates were centrifuged at 14 000 r.p.m. for 30 min and the supernatant was transferred to a fresh tube, and the volume was now increased by 3.5 μ l based on the number of cells pelleted (106) and the average cell volume of 3.5 pl for PC-3 cells (19). Thus, it is calculated that 3.5 μ l \times 3 mM (**1**) = 10 nmol of compound **1** was trapped

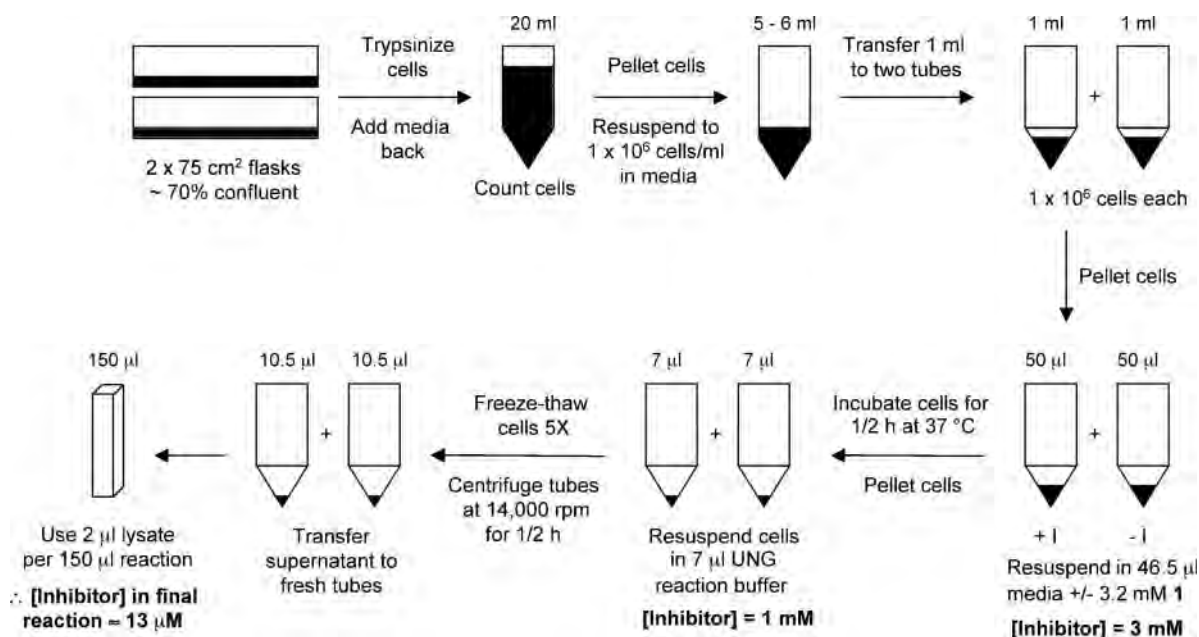


Figure 3. Assay for small molecule inhibition of UNG in cells.

in the pelleted cells. The UNG activity was measured using the PEG-U9 hairpin substrate and 2 µl lysate in a total reaction volume of 150 µl. Assuming equilibration of **1** across the cell membrane, its final concentration in the fluorescence reaction was 13 µM. Initial reaction velocities were calculated using the total change in fluorescence as described (1) and normalized to protein content measured using the BioRad protein assay.

Crystallization of the complex of UNG2 and 3-(2)-A8

Human UNG2 was expressed and purified as described previously (20). A solution of human UNG2 (112.5 µl, 44.2 mg/ml) in a buffer containing 50 mM Tris-OAc, pH 7.0, 150 mM NaCl and 1 mM DTT, was mixed with **1** (12.5 µl, 16.8 mM) in 100 mM Tris-HCl, pH 8.0 and 5% DMSO. The mixture was allowed to incubate at ambient temperature for 30 min, and then centrifuged at 10 000× *g* for 5 min. Co-crystallization conditions were screened using the Nextal PEG Suite library. A total of 300 nl of the complex was mixed with an equal volume of precipitant, and allowed to crystallize at 22°C using the hanging drop method. Crystals were observed within 48 h with 0.2 M potassium thiocyanate, 20% PEG 3350. X-ray diffraction data were collected from a flash frozen crystal in its unmodified mother liquor at the National Synchrotron Light source at Brookhaven National Laboratory (beam line X6A) using a wavelength of 1.1 Å with a ADSC CCD detector Quantum-4. The package HKL2000 (21) was used for data reduction.

The structure was determined by molecular replacement with the program MOLREP using the uncomplexed UNG2 structure (1KHZ) as the searching model. After an initial rigid-body refinement, compound **1** was placed in a difference Fourier electron density. The final model of the UNG2-**1** complex, refined using REFMAC5 (22) with isotropic temperature factors, shows all non-glycine residues in allowed

regions of the Ramachandran plot and excellent stereochemistry (Table 1). Riding hydrogens of protein atoms were used in REFMAC5. The structural statistics are reported in Table 1.

RESULTS AND DISCUSSION

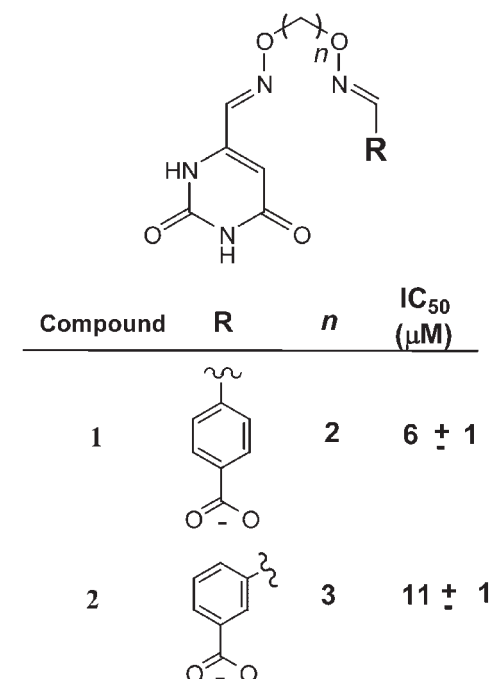
We previously used SFT to identify two uracil-tethered inhibitors of UNG2 from a small aldehyde library containing only 14 different aldehyde binding elements (17). Although these SFT compounds had *K_i* values between 0.3 and 3 µM, the non-uracil binding elements were derived from unstable di- or trihydroxybenzaldehydes. Thus, these initial compounds were prone to air oxidation to the inactive quinone forms, and were not suitable for structural or cell culture studies. We therefore turned to screening a much larger 215 member aldehyde library using our high-throughput molecular beacon fluorescence assay (see Supplementary Data) (17). This screening effort resulted in the identification of two new active mixtures (Figure 4). After deconvolution to identify the linker length that gave rise to inhibition, and purification of the individual inhibitory compounds, the IC₅₀ values were determined (Figure 5A).

The library binding elements that gave rise to the observed inhibition shared a common chemical structure. The two most potent compounds **1** and **2**, with IC₅₀ values of 9 and 11 µM, respectively, both shared formate-substituted benzaldehyde functional groups and short alkyl chain linker lengths of *n* = 2 or 3 (Figure 4). These structure-activity trends suggested the presence of a binding pocket directly adjacent to the uracil binding site that depends on positioning of the negatively charged formate groups of either **1** or **2**. To date, tethering library binding elements to the 6-formyl uracil substrate fragment has brought about increases in binding affinity of up to −3 kcal/mol as compared to the *O*-methyl

Table 1. Crystallographic data collection and refinement statistics

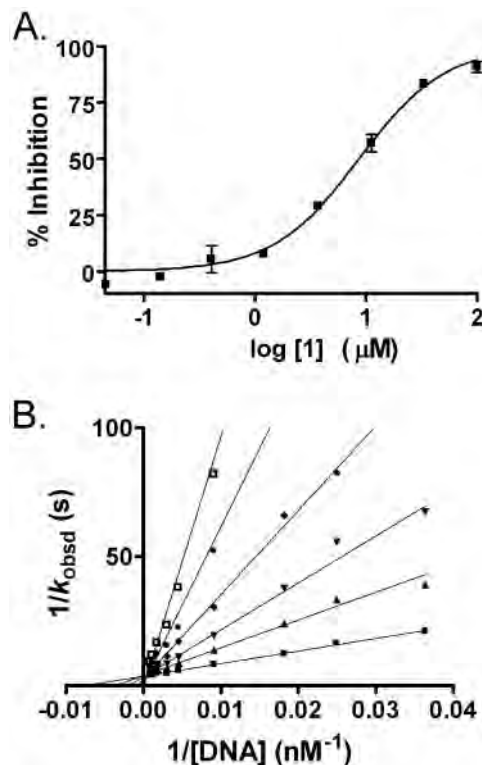
Space group	P212121	
Cell dimensions	$a = 43.2$, $b = 69.1$ and $c = 70.4$ Å	
Resolution range (Å)	49.3–1.3	
R_{sym} (Last shell) ^a	0.06 (0.49)	
Completeness (Last shell)	98.0 (92.4)%	
Multiplicity (Last shell)	4.9 (4.2)	
$I/\sigma(I)$ (Last shell)	10.2 (2.6)	
Number of reflections	51 470	
Refinement		
F Data cutoff in $\sigma(F)$ units	0	
	Number of atoms	Average B (Å ²)
Protein	1808	11.1
Solvent	315	21.0
Ligand	25	13.8
Total	2148	12.6
R -value		0.18
R_{free} (test set of 5%)		0.21
Stereochemical constraints		
Bond length rms (Å)	0.007	
Bond angles rms (degrees)	1.26	
Improper angles rms (degrees)	0.07	

^a $R_{\text{sym}} = \sum_h \sum_j |I_{hj} - \langle I_h \rangle| / \sum_h \sum_j I_{hj}$, where h represents a unique reflection and j means symmetry equivalent indices, I is the observed intensity, and $\langle I \rangle$ is the mean value of I .

**Figure 4.** Structure of inhibitory compounds identified from high-throughput screening. IC₅₀ values were determined for each purified compound and a full mode-of-inhibition analysis and structural characterization was performed for compound **1**. The IC₅₀ value for compound **1** is equivalent to its true K_i .

oxime derived from 6-formyl uracil and methoxyamine ($K_i = 45$ μM) (17).

Since **1** showed the highest activity we investigated its mode of inhibition in greater detail. Interestingly, most of our previously reported SFT inhibitors, as well as the uracil base itself, showed surprisingly complex modes of inhibition with either competitive or partial uncompetitive binding to

**Figure 5.** Inhibition by **1**. (A) Concentration dependence of inhibition. The curve is a nonlinear least-squares fit to Equation 1 ($IC_{50} = 9 \pm 1$ μM). (B) Mode-of-inhibition analysis for compound **1**. Linear competitive inhibition was observed: $K_i = 6 \pm 1$ μM.

two distinct uracil binding sites (17). This surprising complexity, which is not entirely understood at a structural level, was attributed to the presence of a second weak uracil binding site that may be occupied transiently during the process of uracil flipping into the active site. In contrast with the previous complexity, the inhibition patterns for **1** indicated simple competitive inhibition (Figure 5B), with linear Lineweaver–Burk slope replots (data not shown). SFT inhibitor **1** also showed an IC_{50} value against full-length UNG2 that was only 25% greater than the catalytic domain (the full length UNG2 was assayed using cell extracts). The full-length UNG2 protein differs from the catalytic domain by a 90 amino acid N-terminal extension that is involved in nuclear localization and other protein interactions (23). Thus, mode of inhibition analysis indicates that **1** competes for binding to the extrahelical uracil binding site observed in the uracilated-DNA complex shown in Figure 1B and that its inhibitory potency is not affected by the N-terminal extension present in nuclear UNG2.

We also investigated the potency of **1** in cell culture. Since there is no simple marker for assessing UNG2 inhibition in cell culture, an *ex vivo* assay was developed to assess whether the inhibitor enters cells and binds to UNG2. In this assay, cells are treated with a single high concentration of inhibitor (3 mM), and then carefully diluted cell extracts are prepared for fluorometric assay of UNG2 activity (17). Assuming full equilibration of the inhibitor across the cell membrane, and taking into account extract dilutions and measured cell

numbers and volumes (19), the concentration of **1** in the final UNG2 assay was calculated to be 13 μ M (Figure 3). This concentration consistently gave rise to $20 \pm 4\%$ inhibition of UNG2 activity relative to control extracts prepared identically and in parallel, which is $<70\%$ inhibition expected from a competitive inhibitor with a $K_i = 6 \mu$ M. This difference may reflect that (i) **1** is poorly membrane permeable, or (ii) that **1** is not metabolically stable in the intracellular environment. In this respect, oximes are known to be reduced by microsomal NADH cytochrome b5 reductase (24,25).

To evaluate the structural basis for inhibition, UNG2 was cocrystallized with **1** and diffraction data were collected to 1.3 Å resolution and refined to an R_{factor} and R_{free} of 0.19 and 0.22, respectively (PDB ID 2HXM, Figure 6A). In contrast to damaged DNA binding, which leads to a contraction of the active site structure (15) (26), binding of **1** led to only minor structural changes as compared to uncomplexed UNG2 (pdb code 1AKZ), with an r.m.s. deviation over 221 C- α atoms of only 0.58 Å. Despite the differences in induced fit binding as compared to damaged DNA, **1** remarkably shares many of the binding interactions observed in the uracilated-DNA complex (Figure 6B and C). The DNA binding site of UNG2 is composed of a uracil recognition pocket flanked by a deep groove which is predominantly involved in accommodating the single strand of DNA that contains the extrahelical uracil. In the DNA complex, specific hydrogen bonding and aromatic stacking interactions with the extrahelical uracil involving Asn204, His268, Gln144 and Phe158 are observed (Figure 6B). In addition, UNG2 also makes important interactions via neutral and charged hydrogen bonds with the 5' and 3' phosphodiester groups of the deoxyuridine and the 3' phosphodiester group of the 3' adjacent nucleotide using the γ -hydroxyls of Ser169, Ser270 (data not shown) and Ser247, respectively (Figure 6B). The uracil base of **1** shares the uracil interactions seen with the uracilated-DNA complex, with the exception of the catalytically important short hydrogen bond between uracil O2 and His268 (16,27,28). The planar oxime linkage at the uracil side of the tether is observed to extend directly over the space that is occupied by the deoxyribose ring of deoxyuridine in the DNA complex, but then, the alkane linker sharply kinks such that the oxime linkage connecting to the benzylformate moiety nearly perfectly superimposes the path taken by the sugar phosphate backbone of the DNA 3' to the deoxyuridine nucleotide (Figure 6C). This trajectory of the linker presents the carboxylate substituent of the benzyl ring such that it forms a charged tridentate hydrogen bond with the backbone amide groups of Ser247 and Tyr248 and the γ hydroxyl of Ser247 (Figure 6B). These interactions with the carboxylate group mimic those of the 3' phosphodiester group of the nucleotide directly adjacent to deoxyuridine in the DNA complex (Figure 6C). In addition, the oxime oxygen on the uracil side of the tether accepts a hydrogen bond from the γ -hydroxyl of Ser169 thereby mimicking the interaction of the 5'-phosphate of dUrd in the DNA complex. Due to differences in induced fit binding between **1** and uracilated-DNA, the catalytic His268 is too far from uracil O2 to form the strong hydrogen bond seen in the DNA complex. Instead, His268 stacks over the benzyl ring of **1** to form a 3.6 Å π - π aromatic interaction (Figure 6B). Overall, **1** shares three of the four hydrogen bond interactions with the uracil

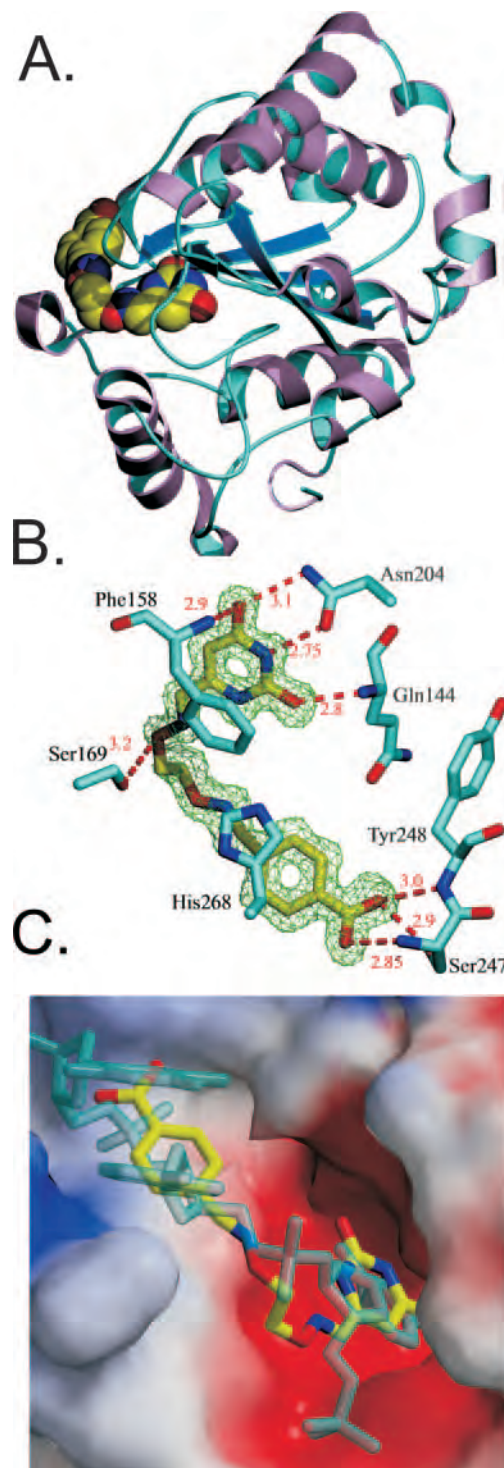


Figure 6. Interactions of **1** and damaged DNA with the active site of UNG2. (A) Global structure of inhibitor-UNG2 complex. (B) Discrete interactions of **1** with the active site of UNG2. The $2F_o - F_c$ electron density map is shown at a contour level of 1σ . (C) Overlay of **1** (gold) with the region of the damaged DNA strand (turquoise) containing uracil and the adjacent two 3' nt (1EMH).

base observed in the DNA structure and three of the five DNA backbone hydrogen bonds.

This structure also provides useful insights into the inhibition provided by compound **2**, as well as our previously

characterized tight-binding SFT compound (**3**, $K_i = 300$ nM) that contains a 2,4,5-trihydroxybenzyl substituent (**17**).

Manual docking studies suggest that the longer three carbon linker of **2** is used to extend its *m*-formate substituent such that it can serve as a DNA phosphate mimic as observed for the *p*-formate substituent of **1**. Our previously reported tight-binding SFT compound contained a *p*-hydroxyl substituent and also a linker that is one carbon longer than **1** suggesting that the *p*-hydroxyl is positioned to form neutral hydrogen bonding interactions in the same pocket occupied by the *p*-formate group of **1**. Apparently, the 20-fold higher affinity of this previous SFT ligand arises from favorable presentation of all of its hydroxyl substituents. It is interesting to note that the 215 member aldehyde SFT library contains only two carboxylate compounds, and both of these were detected as inhibitors in HTS when the correct linker length was employed (i.e. compounds **1** and **2**). Thus, binding elements possessing molecular features similar to the DNA substrate arise more frequently as inhibitors, suggesting that libraries enriched in such motifs might have higher hit rates.

The SFT approach may find general utility in targeting enzymes that recognize extrahelical bases. The flexible alkane tether appears to be an accommodating scaffold that allows favorable presentation of binding elements that are complementary to the DNA binding surface of the enzyme. More generally, substrate fragments (or weak binding ligands) that target enzyme active sites should make excellent starting places for rapid inhibitor development by this or other tethering approaches (29).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank M. Allaire and V. Stojanoff for help during the data collection at the National Synchrotron Light Source. This work was supported by NIH grants GM56834-11 (J.T.S) and GM066895 (L.M.A). D.J.K. was supported by the DOD Breast Cancer Research Program (DAMD17-03-1-1251). Funding to pay the Open Access publication charges for this article was provided by NIH/GMS.

Conflict of interest statement. None declared.

REFERENCES

- Krokan, H.E., Drablos, F. and Slupphaug, G. (2002) Uracil in DNA—occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
- Petersen-Mahrt, S.K., Harris, R.S. and Neuberger, M.S. (2002) AID mutates *E.coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*, **418**, 99–103.
- Lindahl, T. (1974) An *n*-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl Acad. Sci. USA*, **71**, 3649–3653.
- Lindahl, T. and Nyberg, B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405–3410.
- Nilsen, H., Haushalter, K.A., Robins, P., Barnes, D.E., Verdine, G.L. and Lindahl, T. (2001) Excision of deaminated cytosine from the vertebrate genome: role of the smug1 uracil-DNA glycosylase. *EMBO J.*, **20**, 4278–4286.
- Nilsen, H., Rosewell, I., Robins, P., Skjelbred, C.F., Andersen, S., Slupphaug, G., Daly, G., Krokan, H.E., Lindahl, T. and Barnes, D.E. (2000) Uracil-DNA glycosylase (ung)-deficient mice reveal a primary role of the enzyme during DNA replication. *Mol. Cell*, **5**, 1059–1065.
- Stivers, J.T. and Jiang, Y.L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.*, **103**, 2729–2759.
- Seiple, L., Jaruga, P., Dizdaroglu, M. and Stivers, J.T. (2006) Linking uracil base excision repair and 5-fluorouracil toxicity in yeast. *Nucleic Acids Res.*, **34**, 140–151.
- Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T. and Neuberger, M.S. (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in ung-deficient mice. *Curr. Biol.*, **12**, 1748–1755.
- Di Noia, J. and Neuberger, M.S. (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*, **419**, 43–48.
- Nilsen, H., Stamp, G., Andersen, S., Hrivnak, G., Krokan, H.E., Lindahl, T. and Barnes, D.E. (2003) Gene-targeted mice lacking the ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene*, **22**, 5381–5386.
- Priet, S., Sire, J. and Querat, G. (2006) Uracils as a cellular weapon against viruses and mechanisms of viral escape. *Curr. HIV Res.*, **4**, 31–42.
- Malet-Martino, M. and Martino, R. (2002) Clinical studies of three oral prodrugs of 5-fluorouracil (capecitabine, UFT, S-1): A review. *Oncologist*, **7**, 288–323.
- Stivers, J.T. and Drohat, A.C. (2001) Uracil DNA glycosylase: Insights from a master catalyst. *Arch. Biochem. Biophys.*, **396**, 1–9.
- Parikh, S.S., Walcher, G., Jones, G.D., Slupphaug, G., Krokan, H.E., Blackburn, G.M. and Tainer, J.A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083–5088.
- Drohat, A.C. and Stivers, J.T. (2000) *Escherichia coli* uracil DNA glycosylase: NMR characterization of the short hydrogen bond from His187 to uracil O2. *Biochemistry*, **39**, 11865–11875.
- Jiang, Y.L., Krosky, D.J., Seiple, L. and Stivers, J.T. (2005) Uracil-directed ligand tethering: an efficient strategy for uracil DNA glycosylase (ung) inhibitor development. *J. Am. Chem. Soc.*, **127**, 17412–17420.
- Maly, D.J., Choong, I.C. and Ellman, J.A. (2000) Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-src inhibitors. *Proc. Natl Acad. Sci. USA*, **97**, 2419–2424.
- Halgunsset, J., Lamvik, T. and Espevik, T. (1988) Butyrate effects on growth, morphology, and fibronectin production in PC-3 prostatic carcinoma cells. *Prostate*, **12**, 65–77.
- Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N.M., Haug, T., Levine, D.W. and Krokan, H.E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the ung gene and evidence that ung encodes the major uracil-DNA glycosylase. *Biochemistry*, **34**, 128–138.
- Otwinoski, Z. and Minor, W. (1997) Processing of x-ray diffraction data in oscillation mode. *Methods Enzymol.*, **276**, 307–325.
- Ccp4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
- Kavli, B., Sundheim, O., Akbari, M., Otterlei, M., Nilsen, H., Skorpen, F., Aas, P.A., Hagen, L., Krokan, H.E. and Slupphaug, G. (2002) Hung2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hsmug1 as a broad specificity backup. *J. Biol. Chem.*, **277**, 39926–39936.
- Kurian, J.R., Bajad, S.U., Miller, J.L., Chin, N.A. and Trepanier, L.A. (2004) NADH cytochrome b5 reductase and cytochrome b5 catalyze the microsomal reduction of xenobiotic hydroxylamines and amidoximes in humans. *J. Pharmacol. Exp. Ther.*, **311**, 1171–1178.
- Heberling, S., Girreser, U., Wolf, S. and Clement, B. (2006) Oxygen-insensitive enzymatic reduction of oximes to imines. *Biochem. Pharmacol.*, **71**, 354–365.
- Werner, R.M., Jiang, Y.L., Gordley, R.G., Jagadeesh, G.J., Ladner, J.E., Xiao, G., Tordova, M., Gilliland, G.L. and Stivers, J.T. (2000) Stressing-out DNA? The contribution of serine-phosphodiester interactions in catalysis by uracil DNA glycosylase. *Biochemistry*, **39**, 12585–12594.

27. Drohat,A.C., Xiao,G., Tordova,M., Jagadeesh,J., Pankiewicz,K.W., Watanabe,K.A., Gilliland,G.L. and Stivers,J.T. (1999) Heteronuclear NMR and crystallographic studies of wild-type and H187Q *Escherichia coli* uracil DNA glycosylase: electrophilic catalysis of uracil expulsion by a neutral histidine 187. *Biochemistry*, **38**, 11876–11886.
28. Drohat,A.C. and Stivers,J.T. (2000) NMR evidence for an unusually low N1 pK_a for uracil bound to uracil DNA glycosylase: implications for catalysis. *J. Am. Chem. Soc.*, 1840–1841.
29. Erlanson,D.A., Wells,J.A. and Braisted,A.C. (2004) Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 199–223.